UNIVERSITY OF CAPE TOWN

FACULTY OF EDUCATION

THE ANALYSIS AND PREDICTION OF STUDENT PROGRESSION THROUGH DEGREE PROGRAMMES: A COHORT ANALYSIS OF UNDERGRADUATE STUDENTS AT THE UNIVERSITY OF CAPE TOWN

A dissertation presented in partial fulfilment of the requirement for the Degree of

MASTER OF PHILOSOPHY

BY

JANE HENDRY

MARCH 1998

THE ANALYSIS AND PREDICTION OF STUDENT PROGRESSION THROUGH DEGREE PROGRAMMES: A COHORT ANALYSIS OF UNDERGRADUATE STUDENTS AT THE UNIVERSITY OF CAPE TOWN

ABSTRACT

A simplified cohort survival analysis was used to investigate the academic progression of first-time entering undergraduate students within four large bachelors' degree programmes at the University of Cape Town. The rates of graduation, academic exclusion and voluntary drop-out were quantified in relation to the matriculation authorities and prior matriculation performance of the students within each of the four cohorts. The results of the analyses served to identify specific areas of concern with regard to the internal efficiencies in student progression through each of the four degree programmes, and it is suggested that the availability of information of this type will be essential in the attainment of the institutional transformation goals set out in the 1997 White Paper on the transformation of higher education in South Africa.

Significant relationships between the matriculation criteria and the final academic outcomes of students within each cohort were detected using log-near modelling. By means of multiple discriminant analysis, significant predictor variables of the final undergraduate academic outcomes within each cohort were identified. However, the relatively weak discriminatory powers of the multiple discriminant models and the poor predictive accuracy of the associated classification functions suggest the variables included in these analyses did not adequately explain the variability in the final undergraduate academic outcomes of students within the selected cohorts. The extent of the voluntary drop-out phenomenon within each of the cohorts was quantified in relation to matriculation criteria, and further analysis of the cohorts indicated that factors other than academic difficulty appeared to have prompted the greater proportion of the voluntary withdrawals. Those students who had dropped out voluntarily were therefore not included in either the log-linear models or the multiple discriminant analyses.

CONTENTS

APPENDICES

LIST OF EXHIBITS

CHAPTER 1

INTRODUCTION

The University of Cape Town (UCT), is the oldest of the South African universities, and, as such, its history has been closely linked with the changing political dispensations which have held the stage during its existence. South Africa's first universities were intended to serve the white population, although UCT and several other of these institutions gradually opened their doors to students of colour (Grobbelaar, 1995:6)[1]. The 1997 White Paper on the transformation of higher education (Department of Education, 1997) however describes the present higher education system as a fragmented, unequal and inefficient one. The Department of Education has therefore laid down in the 1997 White Paper a broad policy framework outlining a "comprehensive set of initiatives" which aim to redress the inequities of the apartheid system and to transform the South African higher education system into one that is both appropriate to the "new social order" in the country and which is able to respond to national realities, needs and opportunities (Department of Education, 1997:3). UCT and other tertiary institutions are therefore presently addressing the task of formal transformation as laid out in the White Paper on the transformation of higher education.

In outlining the "needs and challenges" of the higher education in South Africa, the White Paper identifies the "...inequitable distribution of access and opportunity for students and staff along the lines race, gender, class and geography" and the "gross discrepancies in the participation rates of students from different population groups" as key concerns driving the imperative for the transformation of the tertiary system (Department of Education, 1997:2). The White Paper presents a vision of a "transformed, democratic, non-racial and non-sexist" higher education system (Department of Education, 1997:4) within which there will be equity of opportunity in terms of both access to higher education programmes and in the opportunity to succeed in these programmes[2] (Department of Education, 1997:5). The Department of Education acknowledges that a core feature of the unitary national higher education system will therefore be the diversification of the "social base" of the tertiary system with regard to

---

[1] Currently, there are 22 universities in South Africa. Ten of these universities were established primarily to serve the white population group, and were categorised by the National Commission on Higher Education (NCHE) as Historically White Universities (or HWU's) (NCHE Report, 1996a:32-33). The NCHE classified the remaining twelve universities as either Historically Black Universities (or HBU's, which were established during the Apartheid era in order to serve South Africa's black population), Distance Education Universities (DEU's) or universities established within the former self-governing states (i.e. Transkei, Venda and Bophutatswana), also primarily to serve the black population.

[2] The term "programmes" is used in this report in terms of the 1996 NCHE report definition of programmes as the "sequential learning activities leading to the award of particular qualifications" (p.84). The NCHE report further distinguishes such "instructional programmes" from the research and continuing education programmes encompassed within higher education (p.85).

race, class, gender and age in particular (Department of Education, 1997:7), and that expansion of the system will be necessary in order to comply with the principles of "equity, redress and development" (p.7) as laid down in the White Paper. Referring to material prepared by the National Commission on Higher Education (NCHE Report, 1996a:64), the White Paper points out that the 1993 participation rate[3] disparity between white students (almost 70% of the 20-24 age cohort) and African students (around 20% of the 20-24 age cohort) is "still substantial, especially when analysed on a programme basis, and by level of qualification", and outlines the Department's commitment to addressing such imbalances via planned expansion of the system (Department of Education, 1997:11). The increased participation of "black students in general, and of African, Coloured and women students in particular" is identified as a major focus of the expansion and equity strategy (p.11). Each tertiary institution will be required to develop a three-year rolling plan within the framework of a national higher education plan; this three-year plan will be required to outline the institutional mission, and to describe its proposed programmes and the predicted enrolments per programme and race and gender equity goals.

On a cautionary note, the White Paper points out that, within the context of the government's macro-economic framework and fiscal policies, it is not yet apparent what actual magnitude of increased overall participation rates, and participation for black students in particular, will be possible (Department of Education, 1997:11). Increased participation could conceivably involve costly physical expansion of tertiary facilities and would almost certainly require the financial support of increased numbers of disadvantaged students. In view of the limitations on real growth in public expenditure on higher education the White Paper acknowledges that tertiary institutions will therefore be compelled to diversify their funding sources, particularly in terms of mobilising additional private funds (p.11). Importantly, concerns with regard to the efficient internal use of funds become implicit in the assertion that progress towards equity of access be paralleled by a concern with "equity of outcomes", i.e. that "increased access must not lead to a "revolving door" syndrome, with high failure and drop-out rates" (Department of Education, 1997:12). The importance attached to equity of outcomes is highlighted in the stated commitment to ensuring that earmarked public funding for equity and redress be linked to "measurable progress toward improving quality and reducing high drop-out and repetition rates" (p.12).

The issue of internal efficiency within the higher education system is dealt with more explicitly within the context of the projected limitations on public expenditure on higher education (Department of Education, 1997:30). The White Paper adopts the view here that there is a great deal of scope for improvement in internal efficiency within higher education

---

[3] These gross participation rates, calculated as the ratio of total student enrolments to total population in the 20-24 age cohort, were determined according to UNESCO norms (Department of Education, 1997:11).

and proposes that both system-wide and internal institutional reforms will be required in order to "reduce wasteful expenditure, improve efficiency and enhance quality" (p.30). Cost reduction measures suggested include open and transparent performance-based public-funding, the rationalisation of institutional programme and service provision and the development of distance education and "resource-based learning" as less labour-intensive teaching and learning methods. Moreover, potential gains in internal efficiency via the improvement of student through-put and completion rates are identified, and it is proposed that progress in this regard be supported by student academic development and support systems, and by "more focused or targeted public funding measures" (Department of Education, 1997:31). In the description of this proposed system of "tightly targeted" public funding (referred to as "goal-orientated, performance-related public funding") to the tertiary education sector, the White Paper again stresses the underlying efficiency objectives which are intended to result in the improvement of student access paralleled by increased student progression and graduation rates. In addition, the new public funding framework, as outlined in the White Paper, will be closely linked with the three-year institutional plan; this will encompass the institution's mission, enrolment targets by programme, race and gender equity goals and measures, etc (Department of Education, 1997:33). Whilst the White Paper advances the view that inclusion of student completion rates within what will become a 'funding grid' encompassing broad fields and levels of study is not appropriate at present, the Department proposes to monitor student progression and graduation rates in the context of institutional progress towards realising performance improvement objectives outlined in its strategic plan (p.33).

In terms of funding for these changes, the White Paper points out that about 85% (on average) of the public funding to universities and technikons is currently "formula funding" (i.e. determined by the subsidy formulae[4]) to these institutions; the non-formula balance of the public funding includes financial aid allocations, funds for capital works, the servicing of loans and other "ad hoc property related charges" (Department of Education, 1997:31). The White Paper stresses that the current 85%:15% split between formula and non-formula funding to universities and technikons cannot address the "urgent needs to lay the foundations of a transformed system", and hence proposes to increase the proportion of earmarked (non-formula) funding in the years 1998-2000 (Department of Education, 1997:32-33). It is envisaged that these earmarked funds will be used to increase funds for student

[4] Formula funding of universities is currently accomplished by means of a complex mechanical formula which takes cognisance of projected student enrolments, student success rates and projected research output. The subsidy total for any university is calculated by subtracting from the formula total the portion which the institution is supposed to generate from its private resources (e.g. tuition fees and investment income), multiplied by an "a-factor", which was introduced in order to "scale down" subsidy payments to universities during a period of unaffordable growth rates in student enrolments in these institutions (NCHE , 1996b:41-44).

financial assistance, and to fund a targeted program of institutional redress particularly with regard to physical facilities (p.33).

The foci emerging from the discussion above are:

- that tertiary institutions will be required to submit rolling plans outlining their population group and gender targets and measures to achieve these targets;
- that the transformation of the tertiary system, which may well necessitate the expansion of student enrolments, will have to take place within national fiscal constraints, such that
- the goal of equity of access, which is complemented by the objective of equity of outcomes, will have to be met "from a strategic mix of funding sources" necessitating increased private contributions, and increased and redistributed "tightly targeted" public sector funding (Department of Education, 1997:31);
- the viability of these goals will require substantial gains in institutional efficiency which are implicit in the stated objective of reducing drop-out and repetition rates, and which are dealt with explicitly in the context of public funding and financial efficiency.

Higher education institutions are therefore confronted with the dual, and possibly conflicting, demands of broadening access whilst simultaneously improving student progression and graduation rates. The historically advantaged institutions (i.e. the HWU's) may well be required to achieve these goals against a backdrop of financial constraints resulting both from national fiscal disciplines and from the redistribution of public funding within the tertiary sector. At the same time, the White Paper points out that the "chronic underfunding", systematic repression and resistance culture within black education during the apartheid years have had detrimental effects on the preparedness of black students for higher education (Department of Education, 1997:12). The under-preparedness of this group, who form the prime focus of the policy of expanded access, will place added financial strains on universities during the process of transformation: clearly the accelerated provision of bridging programmes and the development of comprehensive student support services envisaged by the White Paper in order to ensure equity of outcomes (Department of Education, 1997:13) among this group will have substantial cost implications. Improvements in institutional efficiency are therefore likely to become a major focus within historically advantaged institutions such as UCT, particularly in view of the Department's stated intention to monitor student progression and graduation rates as indicators of institutional performance.

In its synthesis on efficiencies in higher education in developing countries, the World Bank (1994:19-20) has identified the following factors which tend to drive up costs per graduate, and hence to diminish internal efficiency in tertiary systems:

- under-utilisation of facilities
- low student:staff ratios
- low graduation rates

- high repetition and drop-out rates
- devoting a large share of the budget to non-educational expenditures such as student grants and subsidised student services.

While the reversal of these may be desirable in the abstract, because of the emphasis on equity in terms of access to post-secondary institutions it is unlikely that UCT (and other South African universities) will be easily able to follow the recommendations of the World Bank (1994:67) in enrolling only as many students as the University can "responsibly teach" and accepting only those who can derive maximum benefit from their studies. It would appear that South African universities will rather be compelled to expand their enrolments and to improve equity in terms of outcomes within the context of scarce public funds. It will therefore be crucial that institutions strive to function as efficiently as possible in order to obtain the greatest benefit from their limited resources.

There are however a number of further constraints which limit the policy options available to universities. For example, although South African institutions could improve their internal efficiencies by optimising the use of facilities and increasing student:staff ratios, the feasibility of this approach within the context of expanded access to "under-prepared" students (who may require intensive, small group tuition) is questionable. Similarly, in view of the discussion above, it appears unlikely that significant savings on non-educational expenditure (such as student support and counselling services) will be possible. The focus is therefore likely to therefore fall on the gains in institutional efficiency which would result from improved student progression and graduation rates and from reduced student attrition rates. As Bean (1990a:147) points out, the loss of a single bachelors' student during their first year of study equates to the loss of three years (or four, in the case of a four-year bachelor's programme) of potential fee income to a university. This relatively simple view obscures critical factors such as the opportunity costs and burden of possible loan repayments borne by such students, as well as the wastage of scarce institutional resources and, in the South African context, loss of subsidy income under the current formula funding system. Financial pressures will therefore force UCT (and other tertiary institutions) to formally manage its enrolment in order to minimise student drop-out or attrition (thereby maximising retention) and to reduce rates of repetition.

Within the broad South African tertiary system, there has been relatively little research into the academic progress, retention and attrition of students either on an institutional basis, or within particular programs within institutions. The need for a greater institutional focus on internal efficiencies as outlined above also implies that South African institutions could benefit from the availability of predictive models of student retention in the South African context which would enable them to (i) detect potentially successful students amongst groups

of applicants and (ii) to target for remedial intervention those students at risk of academic exclusion.

The research reported in this report thus represents an attempt to use cohort survival analysis in the descriptive investigation of the patterns of student progression and attrition at a selected South African tertiary institution (the University of Cape Town), coupled with an exploration of the use of two types of statistical modelling (namely log-linear modelling and multiple discriminant analysis) in the development of predictive models of undergraduate academic outcomes within selected bachelors' programmes at UCT.

An account of the efficiency-related expediency of conducting this type of research at UCT (see Chapter 3 of this report) follows the overview of the literature on student retention studies presented in Chapter 2. The research methods used are discussed in Chapter 4, and the research findings which are presented in Chapter 5, are discussed in Chapter 6. The log-linear and multiple discriminant models are developed in Appendices 1 and 2 of the report respectively.

CHAPTER 2

STUDENT RETENTION STUDIES: AN OVERVIEW OF THE LITERATURE

There is a rich literature documenting a wide range of student retention studies, which have been conducted largely within the United States and Canada. Many of these studies have focused on the understanding of the phenomenon of student attrition, with the ultimate aim of improving institutional student retention. The discussion below therefore aims to construct a brief overview of the emergent theories of student retention, of the methods commonly used in student retention studies, and of the use of predictive models of student retention in the context of postsecondary education.

2.1 STUDENT RETENTION THEORIES AND THE IDENTIFICATION OF KEY FACTORS IN STUDENT RETENTION

Tinto's (1975) exploration of student attrition in relation to distinct academic and social factors, including initial student characteristics, is widely cited and appears to have greatly influenced research in the area of student retention. The range of theories of student retention, which has evolved from studies of retention in relation to a wide range of variables, was recently summarised by Sarkar (1993:3-4), who compiled a comprehensive categorisation of the theories of student retention which have emerged since publication of the Tinto model. This author has broadly classified student retention theories according to the following categories which are based on the inherent emphasis on particular factors within these theories:

- psychological theories, which focus on specific personality traits and view retention as a function of a student's resultant ability to successfully complete his/her programme of study. Such theories therefore focus on internal forces (i.e. those within the student) in relation to student retention;
- societal theories, which focus on the importance of students' race, social class or on the "barriers and hurdles" which must be overcome en route to successful programme completion;
- economic theories, which stress the importance of the financial situation and the financial aid status of the student in relation to student retention;
- organisational theories, which emphasise the effect of factors such as the structure and size of the institution and staff:student ratios etc on student drop-out rates; and

- interactional theories, which examine the relationship between student expectations of an institution and the associated level of integration into the academic and social environment within their chosen institution.

Theories of student retention could therefore be more broadly classified as those which focus on "internal factors" on the one hand (those within the student, i.e. the psychological theories) and those which focus on factors "external" to the student (i.e. societal, economic and organisational theories of student retention). Interactional theories encompass both internal and external factors which impact on student retention. For example, Sarkar (1993,p.4) points out that a number of studies have indicated that students are more likely to shoulder a greater financial burden in order to stay in college when they view their experience of higher education in a positive light. This could indicate that particular theories of student retention (economic theories, in this case) should not be applied in isolation.

Furthermore, based on a comprehensive study examining the impact of a range of variables relating to students at the Saskatchewan Institute of Applied Science and Technology (SIASST), (including goal commitment, educational ability, academic/social integration, various student characteristics such as race, gender, marital status, and labour market conditions), Sarkar concluded (1993:12) that student retention is a complex function of many variables, both internal and external. This author highlighted the importance of institutional awareness with regard to subgroups of students identified as being at risk of dropping out, particularly against the backdrop of a diverse and changing student body (p.12).

The findings of a more recent study conducted by Sandel and Sydow (1997) at the Mount Empire Community College (MECC) - a member institution of the Virginia Community College System - in order to address the problem of student attrition, support the conclusions reached in the SIASST study described above. In the MECC study, a survey was conducted amongst all students who had withdrawn from classes prior to completion of the fall semester of 1995; this part of the study was complemented by a survey conducted amongst a range of academic and non-academic staff in order to gauge their perceptions of the reasons for student drop-out. The results of the student survey supported the conclusion that a range of factors (including gender, age, work and family responsibilities) were important in student retention. The study also identified a number of personal and academic obstacles to student retention. Whereas the staff survey revealed that "internal influences" (such as low motivation or self-esteem, unrealistic goals and immaturity) were most often perceived to be obstacles to retention, the students surveyed more commonly cited "external factors" (e.g. health, marital and financial factors) as important obstacles to the realisation of their academic goals (Sandel and Sydow, 1997:62-64).

A number of methodological approaches have been employed in studies of the factors affecting student retention, the results of which have been used in the development of the theories of student retention described above. An outline of the range of methods is presented below in order to clarify the methodological basis of the present study.

2.2 RESEARCH METHODS USED IN THE INVESTIGATION OF STUDENT RETENTION

In his synopsis of the common methodological approaches to understanding postsecondary student attrition[5] and progression, Bean (1990b:173-179) has distinguished four types of student retention study, which differ in their approach to data collection. The four types of study described by Bean are (i) "autopsy studies", which aim to collect descriptive data on student dropouts by means of an interview or questionnaire; (ii) "cross-sectional studies", which gather information on a number of students (with regard to e.g. academic performance, social integration, attitudes and intentions) and subsequently track these students in order to identify those who stay in college and those who leave; (iii) "qualitative studies" within which student attrition is explored by means of observation, open-ended interviews with students, academic staff and parents etc; and (iv) "longitudinal or cohort" studies, by means of which data is collected on a particular group or cohort of students (such as those registering at an institution for the first time) at more than one time during their academic careers in order to study their academic progression.

Bean (1990b:174-179) points out that there are disadvantages to the use all of these techniques. For example, autopsy studies, although easy to conduct and to interpret, are the least reliable in that they depend solely upon the potentially biased perceptions of past students (1990b:174). Similarly, Bean points out that the data gathered in cross-sectional studies may be inaccurate since it focuses on the investigation of one set of students at one particular time, so that the attitudes of the group could be distorted by the unique characteristics of the given group of students, or by the impact of an isolated event (the author cites the example of the dismissal of a "popular professor" as such a one-time event) on student attitudes (Bean, 1990b:176). The major drawbacks of qualitative analysis, in the view of Bean (1990b:179), are the time-consuming nature of the data gathering, transcription and analysis involved, and the doubtful generalisability of the ultimate results (which are commonly derived from a small sample of between ten and thirty participants). Bean (1990b:177-178) is therefore of the opinion that longitudinal studies potentially yield the "best quantitative, descriptive and analytical data" for understanding student attrition, but warns that longitudinal studies may become very complex to conduct and to communicate "in a meaningful way".

---

[5] Student attrition includes both voluntary drop-out phenomenon and the institutional exclusion of students on academic grounds.

The technique of "cohort survival analysis", as described by Ewell, Parker and Jones (1988:8-10), is a longitudinal tracking system which is used to measure student persistence over time (i.e. the proportion of students still enrolled at the institution) and to determine graduation rates in relation to selected entry criteria. In cohort survival analysis, entering students are assigned to unique cohorts identified on the basis of their initial enrolment (e.g. entry year and type of degree enrolled for) and are subsequently "tracked" through the institution until their departure as either graduates or drop-outs; these authors stress that complete cohorts, as opposed to samples of cohorts, are generally used in order to "provide credible program-level statistics" (Ewell, Parker and Jones, 1988:10). Unlike the "autopsy", "cross-sectional" and "qualitative" approaches to the investigation of student retention defined by Bean (1990b:173-179), cohort survival analysis therefore provides a means of quantifying graduation rates and "the time to degree", of determining attrition rates, and of identifying the most common points of drop-out during the undergraduate phase, either within the context of the cohort as a whole or in relation to specific groups of students identified within the cohort.

The research reported in this thesis therefore employs the cohort survival approach to distinguishing patterns in student progression and attrition, with a view to identifying the factors involved in student retention in selected undergraduate degree programmes at UCT. Whilst this type of research could potentially provide both a basis for the development of student retention programs and a means of identifying those groups of students at risk of academic exclusion, enrolment management initiatives at South African tertiary institutions such as UCT would also benefit from the availability of predictive models of student retention which could inform admissions policies within particular areas of study. An overview of the studies in the area of prediction of student retention, and of the types of statistical models commonly used in forecasting student success, is presented below.

2.3 PREDICTIVE MODELS OF STUDENT SUCCESS IN HIGHER EDUCATION

Over the past thirty years, an extensive literature has evolved (in parallel with the student retention literature) around predicting performance in higher education. In his review of this literature, Mathiasen (1983:384) examined a wide range of studies which typically explore the relationship between college performance and (i) high school performance, (ii) performance on college entry exams, (iii) study behaviour and attitudes towards study, and (iv) a range of personality traits. This author concluded that successful college students had generally excelled in high school and had performed well in college admissions tests (p.384). Success in college also appeared to be correlated with academic motivation, and achievement orientation, and "good study habits" (Mathiasen, 1983:384).

Many studies of student retention in relation to the predictor variables listed above have been purely descriptive, reporting on, for example, observed differences in student retention or

success rates in relation to key variables. However, a number of studies have explored the possibility of developing statistical models for use in the prediction of student outcomes; this type of predictive capability could be extremely useful in the formulation of admissions policies and in the identification (for remediation purposes) of currently enrolled students who could be considered to be at risk of academic exclusion. These studies have applied methods such as multiple regression analysis, ordinary least squares path analysis and LISREL (linear structural relationships) to retention research. Bean (1990b:179-181) has outlined the application of these predictive methods, concluding that although LISREL has the greatest potential for statistical modelling student attrition, its methodological complexity may detract from its usefulness as an administrative decision-making tool (Bean, 1990b:181). The outline of the application of statistical methods in the prediction of student academic outcomes which follows therefore focuses on the less complex and hence more commonly employed predictive techniques of multiple regression analysis, logistic regression and discriminant analysis.

The technique of multiple regression analysis is frequently used in the development of predictive models where the outcome (dependent) variable is ratio-scaled, for example in Hamilton's (1990) prediction of grade point average (GPA) upon degree completion. Pascarella and Terenzini (1979) employed a stepwise multiple regression in their definitive investigation of the relationship between "student-faculty contact" (i.e. contact between students and academic staff) and student persistence at Syracuse University, in which they concluded that the frequency of such contact was positively related to persistence. Similarly, Whitaker (1987) employed regression analysis in an investigation of the effects of age, marital status, socio-economic status, goal, academic integration, financial concerns and support, social integration and satisfaction with college in student persistence amongst two-year college students; this analysis, which was conducted separately for white and "nonwhite" students, showed that academic integration was the strongest predictor variable amongst both groups, but that the impact of this variable was doubled amongst the "nonwhite" group.

However, in their criticism of the use of multiple regression analysis in student retention studies, Everett and Robins (1991:28) point out that models predicting successful student performance in relation to scores in entry tests should take into account the dual constraints of (i) the monotonically increasing nature of the relationship (in general, a higher test score should predict a greater chance of success) and (ii) the observation that the associated probabilities of success should lie within the range zero to one. These authors therefore warn against the use of linear methods of regression and correlation because of the possibility that the predicted probabilities of success could lie outside of the desired zero to one range. In addition, Everett and Robins (p.28) point out that quadratic or higher polynomial models do not solve the range of probabilities problem and that such models are not inherently isotonic.

These authors therefore suggest (p.28) that logistic regression is more suited to the prediction of student performance.

Logistic regression (or logit) is a non-linear estimation method, and has been employed as a predictive technique where the distribution of the outcome variable is non-linear or even discontinuous (for example a binary variable such as persistence vs non-persistence). For example, Levin and Wyckoff (1988) employed logistic regression in their work on identifying students most likely to succeed in engineering. This analysis revealed that study attitudes (measured in this case as anticipated weekly study hours), gender, prior performance in non-science subjects and the reason for choosing to study engineering all contributed to the prediction of success in engineering. Similarly, Voorhees (1987) used logistic regression in an exploration of persistence in the community college environment which revealed that a range of indicators of academic integration (including GPA, weekly study hours, and frequency of informal interaction with academic staff) were independent of student persistence.

In their more recent exploration of the predictive value of the Tertiary Entrance Score (TES), the Australian Scholastic Aptitude Test (ASAT), gender and field of study amongst students commencing full-time studies at the University of Western Australia (UWA), Everett and Robins (1991) ran a variety of hierarchical logit models testing various combinations of the predictor variables in relation to first-year performance. These authors were able to show that a model including TES and introducing both sex and area of study additively (but not interactively) was sufficient to explain the observed data (Everett and Robins, 1991:31). Furthermore, the predictive capability of the logit analysis enabled the authors to postulate that students with an unacceptably low chance of first-year success were being admitted to study agriculture and economics and, to a lesser extent, to study science and engineering (Everett and Robins, 1991:35). The analyses also suggested that the quantitative component of ASAT alone was comparable to the TES (of which ASAT forms a small component) as a predictor variable. The application of logistic regression in this work therefore served the dual purposes of revealing that admissions policies at the University of Western Australia were not effective in screening out potentially unsuccessful students in certain fields of study, and of establishing that a simplified admissions test could be successfully replace the TES.

Where academic outcomes are classified in terms of categorical variables or group membership (for example, membership of either the group of graduates, the group of drop-outs or of the group of continuing students) discriminant analysis can be employed in the identification of predictors of group membership; the classification functions resulting from discriminant analysis can also be used as predictive functions, i.e. for the identification of likely group membership of new cases (students, in this context) based on the values of the values of the significant predictor variables detected in the analysis. Neely (1977) employed

discriminant analysis in order to detect the best predictors of four categories of college educational outcome at the University of Northern Colorado, *viz.* "college degree within four years", "still enrolled after four years", "not enrolled due to academic suspension" and "left college in good academic standing". In this study, the ACT (American College Testing Program) score and high school percentile rank were found to be the best predictors of college graduation. By means of a stepwise discriminant analysis, Weidman (1985) later identified financial aid status, grade point average (GPA), matriculation status and age as significant predictors of persistence amongst adult (i.e. mature) students at Youngstown State University; these four variables explained 25% of the variance in the study population. The classification functions derived in this analysis correctly classified 81% of the sample population (88% of the persisters and 57% of the non-persisters) in terms of their ultimate educational outcomes.

Where both the academic outcome and the predictor variables are categorical in nature, the relationships between these variables can be explored by means of chi-square analysis, or by means of log-linear analysis where there are more than two grouping variables involved in the study. By means of log-linear analysis, the statistical significance of effects of the potential predictor variables (known as design variables) on the academic outcome variable (known as the response variable), and their of interactions, can be tested. For example, Badenhorst, Foster and Lea (1990) employed log-linear analysis in their investigation of the academic performance of a group of Psychology I in relation to a range of categorical variables - this study is described more fully in section 2.4 below, which reviews South African research into the prediction of students performance.

In summary, despite the wide range of potential predictor variables investigated, the research described above has shown consistently that prior academic performance (during secondary school) and achievement in standardised postsecondary admissions tests are valuable predictors of student success at the postsecondary level. In some cases, indicators of academic integration have been shown to have predictive value, and a range of biographic variables and personality traits have also been tested (with varying results) for predictive capacity in the postsecondary environment. Postsecondary "success" itself has been construed by different authors in terms of either the GPA at the end of the first year of study, or the graduating GPA, or the categorical undergraduate academic outcome. The selection of particular forecasting techniques by different authors clearly relates to both the nature of outcome variable (i.e. a ratio-scaled variable such as the GPA, or a grouping variable categorising the undergraduate academic outcome), and the range of potential predictor variables being examined.

The overview of South African research into the prediction of student success at the university level presented below indicates that there are parallels between local research and

that conducted abroad in terms of the potential predictor variables analysed, and also that innovative methodological approaches to the prediction of student success have been employed in some instances (e.g. Badenhorst *et al*, 1990; Jawitz and Scott, 1997). An important difference between local research and that described above is the absence of a South African national system of standardised postsecondary admissions testing (such as the American ACT and the Australian TES), and hence a greater reliance on performance in the matriculation (school-leaving) examination as an indicator of potential success in the tertiary environment. Coupled with this reliance on matriculation performance as a predictor of tertiary success is the observation that the matriculation examinations within the old, racially segregated secondary system were not "uniformly good predictors" of university performance (Mitchell *et al*, 1997:382). The matriculation authority[6] has therefore become an important variable in the prediction of university performance in relation to matriculation performance within the South African context.

2.4 PREDICTING STUDENT SUCCESS IN THE CONTEXT OF SOUTH AFRICAN HIGHER EDUCATION

What then are the factors which influence student retention at South African universities? Within the context of South African universities, it is clearly becoming increasingly important that student retention studies be carried out institutionally in order to quantify key efficiency-related variables such as rates of graduation and attrition (due to both academic exclusion and voluntary drop-out), and to establish benchmarks indicating, for example, acceptable rates of graduation and academic exclusion within particular faculties or programmes. Should these studies expose particular problems, for example with regard to high rate of academic exclusion or the low rate of graduation amongst certain groups of students, it could be useful to further investigate these problems in terms of the theories of student retention outlined above.

Academic progression within the South African higher education system has been studied far less intensively than that within the American postsecondary system, for example. Moreover, many of the local research efforts have focused on first-year performance (sometimes within a specific course) only, and the South African research in the area of student progression has been characterised by a lack of methodological consistency. Importantly, Smith (1992:10) points out that although there has been local interest in the area of predicting student success at university for some time, much of the South African research has focused on failures amongst white students at different universities.

[6] An explanation of the South African "old" (before 1996) and "new" (from 1996 onwards) matriculation authorities is presented in Chapter 4 of this report.

A comprehensive study commissioned by the Committee of University Principals (CUP) in 1982, for example, suggested that matric aggregate was the strongest single predictor of university success amongst a sample of 6527 (largely white) students (Stoker *et al*, 1985:p.98). Van Wyk and Crawford (1984) subsequently explored the correlation between matric symbols and the marks obtained in a first year ancillary physics course at the University of the Witwatersrand over 5 years, concluding that the weighted matric points[7] rating (i.e. where the points for mathematics and the best science subject within the physical science, biology, geography and physiology group were doubled) was a useful predictor of success (Van Wyk and Crawford, 1984:9). However, these authors did not explore the possible effects of either race or matriculation authority in the relationship between university performance and matriculation results in their research.

Subsequently Badenhorst, Foster and Lea (1990) examined the results of almost two thousand Psychology I students at UCT by means of log-linear analysis, and reported that matric aggregate and three explicit demographic variables (home language, population group and matriculation authority) were important predictors of success in this course. These authors conceded that there were relatively few African students within the study group, but were of the opinion that the dominance of the matric aggregate variables was of particular importance in the determination of admissions policies (Badenhorst *et al*, 1990:44). Badenhorst *et al* (1990:44) pointed out that the log-linear analysis did not detect significant interactions between the predicator variables, indicating that the matric aggregate was therefore an "equally good predictor (of the level of achievement in Psychology 1) at every level of Population, Language or Matric Authority".

More recently, Jawitz and Scott (1997) reported on their investigation into the effects of a range of factors (including population group, gender, matric points and participation in ASPECT, the Academic Support Program for Engineering in Cape Town) on the retention of first-year engineering students at UCT. In this study, retention rates were calculated as follows :

$$RR = \frac{\text{(no. of engineering graduates) + (no. of students continuing in engineering)}}{\text{no. of students who registered for first year engineering at UCT}}$$

By means of analysis of co-variance (where the matric points rating was used as a co-variate "because of the strong association between matric points and performance"), Jawitz and Scott (1997:5) detected significant differences in retention rates with regard to both population group and gender; the engineering discipline was also found to significantly affect the retention rate (Jawitz and Scott, 1997:7). Jawitz and Scott (1997:8) interpreted their

[7] The calculation of South African matric points and faculty points is explained fully in Chapter 4.

findings to suggest that "elements associated with particular disciplines, such as curriculum, standards of assessment and the teaching and learning environment contribute significantly to the (engineering) retention rate". In their concluding comments, these authors recommend the establishment of a uniform system for monitoring efficiency within the tertiary education system, which they believe would "provide valuable input to on-going academic development initiatives in engineering education in South Africa" (Jawitz and Scott, 1997:8).

In an investigation spanning the twelve year period between 1980 and 1992, Mitchell, Fridjhon and Haupt (1997) analysed the first-year performance of 29 254 students admitted to the University of the Witwatersrand in the year immediately following their matriculation year. These authors attempted to correlate first-year performance (as measured by the average first-year mark) and matric performance (as measured by the matric average, or mean matric mark). The results of this investigation support findings that the matriculation average is a reasonably good predictor of first-year university success, but indicate that the accuracy of prediction varies according to the matriculation "type" (or authority), the school of origin and the registration faculty within the University. Students who had written the examinations of the Indian Senior Certificate, the Coloured Senior Certificate and the Department of Education and Training examinations appeared to be at a substantially greater risk of first-year failure than those students who had matriculated within one of the white provincial matriculation authorities or the JMB/IEB (Joint Matriculation Board or Independent Examination Board). Although these authors acknowledged that they had not anticipated a high level of correlation between matriculation and first-year university performance given the marked difference between the school and the postsecondary learning environments, they hypothesised that if the different matriculation examinations prepare students equally well for their first year of university, the coefficients of correlation between matric average and first-year average should be similar for all matriculation authorities. This hypothesis was rejected, hence the suggestion that "there are so many factors which underlie success at university that overemphasis on matriculation performance is unwarranted" (Mitchell *et al*, 1997:387). Based on their findings, Mitchell *et al* (1997:387) argue that there is a strong case for the establishment of a separate (i.e. independent of the school leaving examinations) university entrance examination.

Mitchell *et al* (1997:382) observe that South African university students come from a "wide, often impoverished, and non-uniform knowledge substrate". This perception is confirmed by O' Connell (1994:173) who describes how the old racially segregated education system has severely hampered the academic development of black students (i.e. African, coloured and Indian students). This author maintains that African students in particular face substantial barriers to "the acquisition of tertiary level knowledge and skills" (O' Connell, 1994:173). The discussion presented in this section however indicates that the academic progression of the various groups of students within South African universities has not been thoroughly

researched, and that there has been a lack of methodological consistency within the local research, both of which serve to compromise the generalisability of the findings of the work that has been done.

In view of the paucity of descriptive and predictive student retention research in South Africa, and of the Ministry of Education's stated goals of equity of access and equity of outcomes within the transformation of the tertiary system (see Chapter 1 above), it would appear that there is a pressing requirement for further investigation into the effects of secondary schooling (as measured by achievement in the various matriculation examinations) on tertiary achievement. The findings of such research will be fundamental to both the development of tertiary admissions criteria, and to the formulation of policies of student support which aim to ensure that all students within an institution enjoy an equal opportunity of succeeding in the tertiary environment.

The possible impact of the transformation initiatives described in Chapter 1 on university finances and on internal efficiencies is discussed in the following chapter (Chapter 3) and is explored using the University of Cape Town as a case study. This chapter aims to demonstrate the necessity of conducting research into student progression in the light of recent trends in student output at UCT, and in view of the financial implications of these trends.

CHAPTER 3

TRANSFORMATION AT THE UNIVERSITY OF CAPE TOWN: SOME INTERNAL EFFICIENCY IMPLICATIONS

The discussion of the White Paper dealing with the transformation of higher education presented in Chapter 1 described how, as a tertiary institution, the University of Cape Town will be required to formulate three-year rolling institutional plans. Within these plans, the University will have to set race and gender equity enrolment targets, not only in general terms but also at the programme level, which could necessitate increasing student enrolments. At the same time, the University will need to establish target student progression and graduation rates and to outline measures put in place in order to achieve these targets. Negotiations around these targets and the University's progress towards them will be fundamental in the government's annual allocation of a block operating grant in support of teaching and related costs. Because of the national policy of fiscal discipline and the proposed changes in the proportions of formula to non-formula public funds allocations within the tertiary sector (see Chapter 1), UCT may well have to address these issues against a background of diminished resources, a situation which will reinforce the need for concerted efforts to identify possible means of improving institutional internal efficiency. An examination of some gross financial and student throughput indicators at UCT in recent years (see below) suggests that there was cause for concern with regard to internal efficiency even prior to the difficulties which could result from the proposed changes to the distribution of government funds within the tertiary sector.

In this chapter, some key financial indicators during the 1988 - 1995 period are discussed, and these are subsequently linked to student statistics for UCT during the same period. In this way, the financial constraints within which the University has been compelled to operate during recent years is related to institutional student throughput, highlighting the necessity for policy and planning in relation to these issues.

Exhibit 3.1 tabulates the changes in UCT's fund income during the 1988 - 1995 period, as a function of funds source. This data is presented as Rands, as real 1988 Rands (deflated according to the average annual inflation rates during this period) and as nominal Rands on the base of 1988 = 100. Clearly government appropriations[8] to the University have declined in real terms during the 1988 - 1995 period (from 90.4 million Rands to 86.0 million Rands). Whilst this decline was balanced to a limited extent by increases in tuition fee income

[8] Government appropriations include total payments under the subsidy formula, government contributions to interest and redemption payments, and government contributions to local authority rates payments.

(especially in the 1988 - 1993 period), the overall effect has been the net 4,01% reduction in real terms of UCT's total income from 234.4 million Rand in 1988 to 225.0 million Rand in 1995.

Exhibit 3.1 : Funds available (Rands millions) 1988 - 1995

	Government Appropriations			Student tuition fees			All other income			Total		
	1988	1993	1995	1988	1993	1995	1988	1993	1995	1988	1993	1995
Rands millions	90.4	191.4	236.1	33.4	89	94	110.6	215.8	287.8	234.4	496.2	617.9
1988 Rands millions	90.4	93.1	86.0	33.4	43.3	34.2	110.6	104.9	104.8	234.4	241.2	225.0
Nominal Rands (1988=100)	100	212	261	100	266	281	100	195	260	100	212	264

Source : The 1988 and 1993 figures are from Bunting (1997:39); the 1995 figures were calculated from the 1995 SAPSE[9] tables.

Against this backdrop of real decreases in funds available, there has been a real increase in current[10] expenditure at UCT. Exhibit 3.2 indicates that real current expenditure increased from 170,1 million Rands to 178,2 million Rands between 1988 and 1995. The observation that the total expenditure figure dropped from 194.0 million Rands in 1993 to 178.2 million Rands in 1995 may be indicative of institutional efforts to curb expenditure during this period. Exhibit 3.2 however shows that there was a small overall increase (0.01%) in total expenditure between 1988 and 1995, but that total expenditure was contained by means of a real decrease in fixed asset and debt service expenditure (from 32.9 to 26.7 million Rands) during this period. This would indicate that UCT has been unable (in real terms) to sustain its 1988 level of expenditure on fixed assets, which includes expenditure on land and buildings, computers and other equipment, furnishings, motor vehicles and library materials (Bunting, 1997:40).

[9] SAPSE, which is an acronym for the South African Postsecondary System, is a system for the annual collection and collation of student and staffing statistics, as well as financial data, for postsecondary institutions. The data required is returned by the institutions to the Department of Education on an annual basis, where it is verified and collated to produce a compilation of "Financial and Related Statements" for each institution, each year.

[10] Current expenditure includes expenditure on staff, on supplies and services, and on student bursaries and scholarships.

Exhibit 3.2 : Summary of total expenditure 1988 - 1995

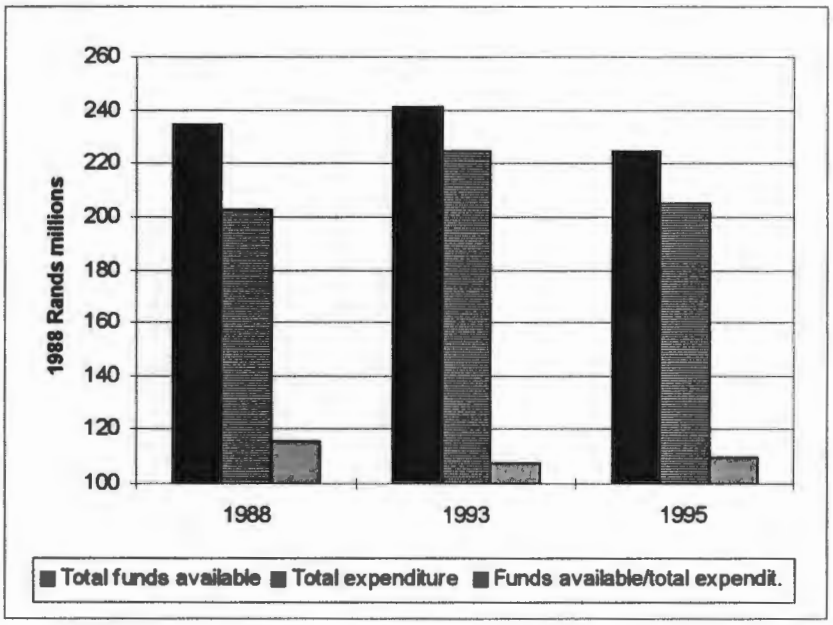
	All current			Fixed asset & debt service			Total		
	1988	1993	1995	1988	1993	1995	1988	1993	1995
Rands millions	170.1	399	489.5	32.9	63.1	73.2	203	462.1	562.7
1988 Rands millions	170.1	194.0	178.2	32.9	30.7	26.7	203.0	224.7	204.9
Nominal Rands (1988=100)	100	235	288	100	192	222	100	228	277

Source : The 1988 and 1993 figures are from Bunting (1997:39); the 1995 figures were calculated from the 1995 SAPSE tables.

The effects of the overall increase in expenditure in relation to the overall decrease in real funds available are shown in Exhibit 3.3 below. Exhibit 3.3 indicates how the real funds surplus (i.e. the difference between total funds available and total expenditure) has diminished during the 1988 - 1995 period. The ratio of funds available to expenditure has declined from 115,5 in 1988 to 109.8 in 1995. The apparent institutional funds "surplus" should not, however, be regarded as a profit since an apparent funds surplus may simply reflect unspent balances of funds already designated for specific purposes (Bunting, 1997:40). What is clear is that there has been an increase in UCT's total expenditure as a function of total funds available, and hence a reduction in the proportion of available funds retained at the end of 1995, in comparison with 1988. Significantly, this situation would appear to have resulted from the limited possibilities for balancing the real decline in government funding with increased tuition fee income and income from other sources[11], as well as from the limits within which institutional expenditure can be curtailed without shedding teaching departments (and therefore, staff). Furthermore, the reduction in the proportion of available funds retained in the latter part of the 1998 - 1995 period implies a reduction in the funds available for investment in the University's long term holdings, which would in turn limit the possibilities for deriving income from investments.

[11] The NCHE report points out that (in global terms) there is already a high level of diversification of income sources within South Africa's universities and technikons, and stresses that the South African tertiary system could only increase its total income from commerce, industry and donation "if the historically black institutions were able to increase their income from these sources to about the same levels achieved by historically white institutions (NCHE, 1996b:13). The NCHE report (p.13) also describes how serious student unrest could result from increases in tuition fees "at rates greater than they are now in the absence of a viable national student financial aid scheme". EduSource report that the costs of student unrest at tertiary institutions related to the issue of unpaid student fees during February and March of 1997 amounted to R4,9 million (EduSource Data News No. 18, p.7)

Exhibit 3.3 : Funds available, total expenditure and income:expenditure ratios 1988 - 1995



Against the background of the decline in available income and concomitant small increase in total expenditure, Exhibit 3.4 shows that there have been marked shifts in student enrolments. Exhibit 3.4 below summarises the shifts in headcount student enrolments and in the total number of graduates during the 1988 - 1995 period:

Exhibit 3.4 : Student output indicators 1988 - 1995

	1988	1993	1995
Headcount enrolments	13280	14200	14908
Total graduates	3244	3795	3707
Graduates as proportion of headcount enrolments	24%	27%	25%
Undergrad. throughput rates	21%	27%	21%

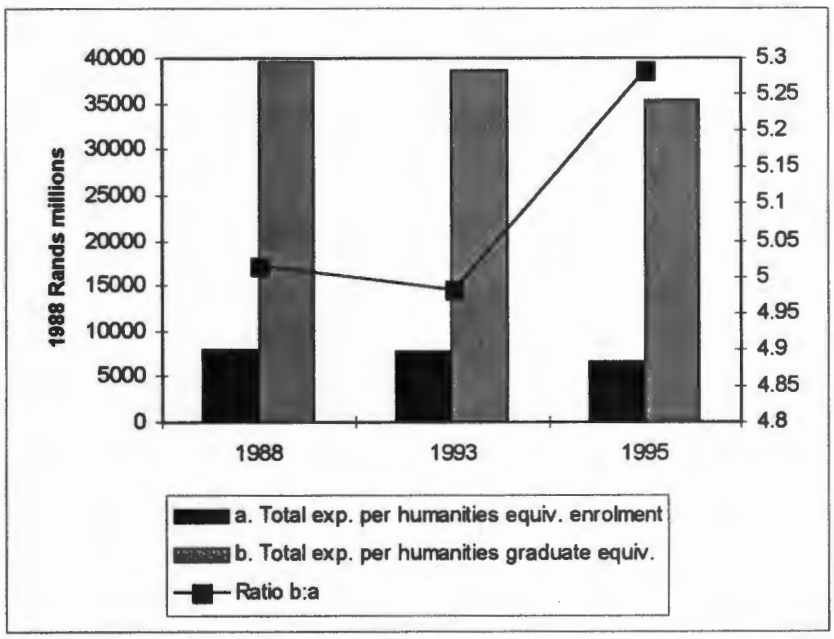
Source : The 1988 and 1993 figures are from Bunting (1997:5-18) and the 1995 figures were calculated from the 1995 SAPSE tables.

The data presented in Exhibit 3.4 show that there was a substantial (12%) increase in student headcount enrolments (from 13280 to 14908) during this period, but also that there was a decrease in the number of graduates between 1993 (3795 graduates) and 1995 (3707 graduates). Although the proportion of graduates (expressed in relation to total headcount enrolments) increased by one percentage point between 1988 and 1995, this proportion dropped by two percentage points between 1993 and 1995. At the same time, the 3-year

undergraduate throughput rate[12] declined sharply from 27% in 1993 to 21% in 1995. These indicators of student output therefore suggest that there was a decline in internal efficiency in terms of both the graduation rate and the rate of repetition and drop-out during the 1993 - 1995 period.

Some key unit cost effects of these dynamics in student enrolments and graduation rates in relation to the changes in total expenditure are depicted in Exhibit 3.5 below. Of particular interest is the change in the relationship between total expenditure per humanities equivalent enrolment[13] and per humanities graduate equivalent[14]. This ratio increased sharply between 1993 and 1995, indicating that there was a marked, (real) relative increase in the total expenditure per humanities graduate equivalent.

Exhibit 3.5 : Total expenditure (Rands millions) (a) per humanities equivalent enrolment and (b) per humanities graduate equivalent, and the ratio of b:a



Source: The 1988 and 1993 figures are from Bunting (1997) and the 1995 figures were calculated from the 1995 SAPSE tables.

[12] The undergraduate throughput rate was calculated for each year by dividing the total number of students completing 3-year bachelors' degrees and diplomas by the total headcount enrolments in these qualifications.

[13] Humanities student equivalents, which are a basic unit of analysis for costing purposes, are calculated by multiplying the full-time equivalent science and technology enrolments by a factor of 2.67 and adding this total to the full-time equivalent enrolments in the broad humanities. The concept of the humanities student equivalent evolved from the current subsidy formula which applies a weighting of 2,67 to each science/technology student for cost purposes, in recognition of the relatively higher costs involved in teaching and research in this area (Bunting (1997:8)).

[14] Humanities graduate equivalents are calculated by multiplying the science/technology graduate total by 2,67 and adding this figure to the broad humanities graduate total.

In summary, UCT experienced a real decrease in government funding between 1988 and 1995, a decrease which was not balanced by increases income from tuition fees and other sources so that there was a real decrease in the total funds available during the 1988 - 1995 period. Although the real increase in total expenditure during this period was very small, the ratio of income to expenditure had declined markedly. During the same period, total headcount enrolments at the University expanded by approximately 12%, so that real expenditure per enrolled student unit declined, as did real expenditure per graduate unit. Examination of the ratio of expenditure per enrolled student unit to expenditure per graduate unit however revealed a sharp increase between 1993 and 1995, indicating that it was becoming relatively more expensive to produce graduates in the mid-nineties.

The discussion in Chapter 1 of this report described how the transformation initiatives outlined in the White Paper on higher education requires UCT to make discernible progress towards broadening student access and simultaneously improving student progression and graduation rates, both of which are to be accomplished within the financial constraints arising from national fiscal policies and from the redistribution of public funding to the tertiary sector. The review of financial and student output indicators presented above however suggests that, contrary to that stated institutional goals set out in the White Paper, the process of producing graduates at UCT is becoming less efficient in financial terms.

It is therefore imperative that UCT identifies means of improving internal efficiency in terms of student progression and graduation rates, and thus it is crucial the University first clearly identifies current areas of *inefficiency*. As has been discussed in Chapter 1, internal efficiency of this nature results from high rates of drop-out and/or academic exclusion as well as from high rates of repetition, which exert a downward pressure on the graduation rate and an upward pressure on the time to degree. The cohort survival analysis described in this report therefore aims to quantify the rates of graduation, repetition and drop-out within the University's four large general academic bachelors' degree programmes in order to establish if and where such internal inefficiencies exist. In addition, this research explores the development of predictive models of student retention within the selected degree programmes with a view to enabling enrolment planners to improve student progression through these programmes by means of identifying potentially successful students, and detecting those likely to require remedial intervention in order to succeed at UCT. The research aims outlined here are further elucidated in Chapter 4.

CHAPTER 4

RESEARCH AIMS, CONCEPTUALISATION OF RESEARCH AND METHODS EMPLOYED

4.1 RESEARCH AIMS

Within the broad South African tertiary system, relatively little is known about student attrition rates and academic progression (in terms of graduation rates or "time to degree"). The annual SAPSE returns, which constitute the only official collection of national higher education statistics, require institutions to report on the overall first-time entering undergraduate attrition (in terms of academic exclusion and voluntary drop-out), and on the number of students leaving each institution without completing a degree/diploma. These attrition data are reported as a function of entrance category, population group and gender, but there is no quantification of this phenomenon in the different degree/diploma programmes offered by institutions. Nor is there any attempt within the SAPSE system to quantify rates of repetition and time to degree even at the level of the total institution, or to locate these within the classification of subject areas used in the system. In other words, the SAPSE data have limited usefulness and applicability in the investigation of institutional student progression and retention, both of which have been identified above as key variables in terms of both equity of outcomes and internal efficiency goals.

By means of a cohort study approach which offers a means of obtaining the necessary data on the longitudinal progress of students enrolled at particular institutions (or within specific programs within an institution), the research described in this report aims to explore the internal efficiencies in student progression through the four large 3-year bachelors' degree programmes at UCT. These are the Bachelor of Arts (BA degree) in the Faculty of Arts, the Bachelor of Commerce (BCom degree) in the Faculty of Commerce, the Bachelor of Science (BSc degree) in the Faculty of Science, and a combination of the BA and Bachelor of Social Science (BSocSc degree) programmes within the Faculty of Social Science and Humanities. This evaluation makes use of a purely quantitative cohort tracking strategy in order to answer important questions with regard to student graduation rates and student retention within the selected programmes, and subsequently attempts to derive predictive models of student progression through these programmes in relation to key matriculation criteria.

4.2 CONCEPTUALISATION OF RESEARCH STRATEGY

The research strategy was conceptualised as a simplified variant of cohort survival analysis[15]. Cohort survival analysis as defined by Ewell, Parker and Jones (1988:5) permits

> (a) [the] construction and estimation of a comprehensive longitudinal picture of student progress that reflects the manner in which students of different kinds move into, through, and out of an institution, and (b) [the] identification of a minimum number of distinct behavioural groups that together constitute the bulk of an institution's enrolment.

The simplified cohort survival analysis used in this research focused on the examination of the first and last undergraduate academic records of students in order to detect graduation rates (and time to degree), rates of academic exclusion and the duration of studies prior to exclusion, and rates of voluntary drop-out and the duration of studies prior to drop-out.

The longitudinal databases contained information on South African matriculants (and those with South African matriculation exemption) who were first-time entering (FU) students, entering a tertiary institution for the first time in the initial year of the study (1992). The cohorts examined therefore consisted of groups of students who were homogenous in terms of their lack of prior experience of tertiary level studies. By contrast, simply tracking the academic progress of a group of first year students in any particular programme (i.e. students in the first year of study for a particular programme) would yield aggregated information on FU students, on those repeating their first year of study in that particular programme, on students transferring into the particular faculty at the first-year level, and on those transferring to UCT at the first-year level from other tertiary institutions; the selection of FU students only effectively removes from the analysis a range of variables pertaining to prior postsecondary experience which could profoundly affect the variability within the aggregated longitudinal data.

The starting point of the analysis, i.e. the 1992 calendar year, was selected in order to allow for the collection of five years of data, spanning the maximum period of enrolment for a three-year bachelors' degree. Student progression was examined in relation to the key matriculation criteria described in the previous chapter, i.e. matriculation authority and weighted matric points[16], which are known as faculty points, and which form the basis of the

---

[15] Cohort survival analysis is not, in itself, a statistical survival analysis method, although the data generated by the analysis are suited to subsequent statistical analysis.

[16] UCT (like all South African universities) uses a Swedish points system to quantify a student's performance in a South African school leaving examination. In the calculation, the student's six best matric symbols are given the following points equivalents : for subjects written at the Higher Grade, an A symbol is allocated 8 points, a B symbol 7 points, a C symbol 6 points and so on; for subjects written at the Standard Grade, an A symbol is allocated 6 points, a B symbol 5 points, a C symbol 4 points, and so on. The points equivalents are then totalled to yield the matric score. Within certain faculties, a weighted matric score (known as the faculty score) is used to assess students' eligibility for admission. In the Science Faculty, the scores for mathematics and the best science subject are doubled whilst in the Commerce Faculty, the scores for mathematics and English are doubled in the calculation of the weighted matric points, or faculty points. Within the Arts and Social Science faculties, no weighting is used, and thus the faculty points are the same as the matric points for any particular student.

admission rating system. The 1992 cohort of students were products of the old, racially segregated secondary schooling system. This system comprised the four white provincial authorities and the JMB (Joint Matriculation Board), the Department of Education and Training (DET) examinations for African students, the Coloured Senior Certificate (CSC) at the so-called "coloured" schools and the Indian Senior certificate (ISC) which was written at schools serving the Indian community. The South African matriculation authorities are currently organised along provincial lines with all students within a provincial public school system writing the same provincial examination. Investigations based on the "old" matriculation authorities could be considered to be outdated in terms of the new organisation of secondary education, but Mitchell et al (1997:382) believe that conclusions drawn from such investigations remain generally valid and "will continue for many years to apply to the schools which were part of the old classification". In other words, whilst the matriculation examination authorities have changed, the schools within which pupils are prepared for these examinations, and which therefore shape their pre-postsecondary experience, will take considerably longer to do so.

4.3 RESEARCH METHODS

4.3a Computer program and data manipulation

Although the data required for the envisaged cohort analysis exists within UCT's on-line student admissions and registration computer system (Heritage), this data is not organised in a format suitable for direct use in a longitudinal database. The starting point in the establishment of the longitudinal database was therefore the conceptualisation of a data format which would permit the type of longitudinal analysis envisaged, and the commissioning of the writing of a computer program which would download the data in the required format from the Heritage system.

Enrolment data pertaining to first-time entering (FU) students registering in 1992 for general academic bachelor's programmes in 4 faculties at UCT (Arts, Commerce, Science and Social Science) were downloaded from the University's on-line student admissions and registration system by means of the customised download program[17]. At the end of the 1996 academic year, the download program was run separately for each of the 4 faculties in the study in order to retrieve (by student number) the following information for each student :

- general biographic data : full name and title, date of birth, gender, population group, entrance category in 1992, matriculation authority, UCT faculty points;

---

[17] Michelle Murry of UCT's Administrative Computing Services was commissioned to write the download program.

- academic progress at UCT : abbreviated, four-field academic records listing the enrolment degree, academic year of study, calendar year and academic progress code for 5 years, beginning with the 1992 academic year.

The downloads were imported from text format into MS Excel spreadsheets, and the sub-set of first-time entering (those who had never before been enrolled at a postsecondary institution) 3-year bachelors' students was extracted for further manipulation and analysis. Only those students who had written the matriculation examinations of one of the current South African matric authorities or who had obtained South African matriculation exemption were included in the database. In this way, data on FU students entering the BA program in the Faculty of Arts, the BCom program in the Faculty of Commerce, the BSc program in the Faculty of Science, and both BA and BSocSc programs in the Faculty of Social Science and Humanities were retrieved within 4 separate working databases.

Because this study employed a simplified cohort survival analysis which focused on the ultimate undergraduate outcomes of FU students, the first phase of data manipulation for each faculty involved the extraction of each student's general biographic data and their first and last undergraduate academic records[18]. In order to achieve this, all intermediate academic records were deleted from the working database, leaving the first and last undergraduate academic records only. The deletion of those students who transferred out of their faculty of origin was also carried out prior to analysis of the cohorts. This step was necessary because students transferring out their faculty of origin could be considered to be on a new academic trajectory governed by the faculty rules specific to their new faculty, and their academic progress was therefore no longer comparable with that of those who remained within their faculty of origin. For example, Science students who transferred to the Faculty of Medicine or to the Faculty of Engineering would have changed from a general academic to a professional degree track with a different minimum completion time and a different set of rules governing academic progression, readmission and academic exclusion. The four fields in the last undergraduate record for each student were used to identify those transferring to other faculties, and the timing of transfer for these students was determined by examining their academic progress within the full cohort download for each cohort (i.e. where each academic record from initial enrolment to the last undergraduate academic record was abbreviated). Once the timing of transfer out of the faculty of origin had been determined, the academic progress of students transferring to other faculties was not analysed any further.

[18] In terms of this study, the last undergraduate academic record was defined as that pertaining to the last year of enrolment at UCT at the undergraduate level, regardless of the academic outcome of this year of enrolment.

Formulae were subsequently inserted into each of the working databases in order to calculate each student's age (at the June 1992 University census, as specified in terms of the SAPSE system) and the duration of their undergraduate enrolment at UCT, which was calculated as the difference between the entry year and the year of the last undergraduate record. Additional data pertaining to the matric English and matric mathematics symbols of individual students were subsequently extracted by student number from the Student Performance Database (an MS Access database which is currently located on UCT's Social Science fileserver) and appended to the working database.

The working database for each of the four cohorts was then used to explore the graduation rates, rates of academic exclusion, rates of drop-out in good academic standing, and duration of undergraduate studies amongst those students who remained within their faculty of origin. These variables were investigated as a function of key indicators of matriculation performance (mainly UCT faculty points and matriculation authority) in order to assess the relationships between prior performance in the matriculation examinations and achievement at UCT. In these analyses, faculty points were grouped within pentile (five point) ranges and similar matriculation authorities were grouped together: DET plus Transkei matriculants were handled as one group , CSC matriculants as a separate group, and those students who had written any other South African matriculation examination as a third group. The grouping of similar matriculation authorities was carried out in order to categorise the matriculation authority-related heterogeneity within the data, which would otherwise have contributed to the overall variability within the statistical models which was not accounted for by the variables included within these analyses. Because two of the faculties (Science and Commerce) employed a weighting system in the calculation of faculty points, the faculty points groupings differed according to the faculty under consideration. However, in all cases the same matriculation authority groupings were used: students who had written either the examinations of the former Department of Education and Training (DET) or the Transkei Matric examinations were referred to as "DET" students; those who had written the Coloured Senior Certificate (CSC) were referred to as "CO" students, and those who had matriculated within any of the other South African matriculation authorities were referred to as "OMA" (Other Matriculation Authority) students[19].

Those students whose last undergraduate academic records reflected the academic progress codes "QUA" or "INQ" were assumed to have completed their undergraduate qualifications successfully, whereas those whose last academic progress code was that of "REN" had been excluded from the University on academic grounds. Those students whose last academic

[19] It was not possible to distinguish JMB matriculants from those who had obtained JMB matriculation exemption within the cohort downloads. Both types of student were therefore included within the OMA matriculation authority group. Students with JMB exemption only would have been amongst those with unknown faculty points (since faculty points are calculated on South African matriculation symbols), and were therefore invalid cases which were not included in the statistical analyses.

records showed an academic progress code other than "QUA", "INQ" or "REN", and were therefore eligible for readmission as undergraduates, were considered to have dropped out of UCT in good academic standing. An explanation of all academic progress codes used is summarised in Exhibit 4.1 below:

Exhibit 4.1 : An explanation of the academic progress codes used at UCT

Academic progress code	Explanation of code
"QUA" or "INQ"	Qualified for degree or diploma
"CON"	Satisfied standard readmission conditions
"COI"	Satisfied special readmission conditions for disadvantaged students
"RET" or "REP"	Did not satisfy either standard or special readmission conditions but received Faculty or Senate permission to re-register
"REN"	Refused readmission on academic grounds

4.3b Statistical analysis and development of predictive models of first-time entering student success at UCT

The Results chapter (Chapter 5 of this report) contains a detailed descriptive analysis of the academic progress of the four selected cohorts in relation to the key matriculation variables described above (matriculation authority and faculty points). The descriptive analyses have been complemented by the application of two types of statistical modelling techniques.

The first of these, namely log-linear modelling, aims to explore the statistical significance of the relationships between academic progress, matriculation authority and faculty points seen in the descriptive Results section. The second statistical method, namely multiple discriminant analysis was employed to serve two purposes : (i) to select from a broader group of matriculation criteria key variables which could be regarded as predictors of the undergraduate academic outcome, and (ii) to produce statistical models predicting undergraduate academic success based on matriculation criteria. The methods of log-linear modelling and of multiple discriminant analysis are described in sections (i) and (ii) below.

It is envisaged that the statistically validated findings of this research could be usefully employed in the setting of admissions criteria with regard to the four degree programmes evaluated here. A second purpose of these analyses could be, rather than restricting access to these programs on the grounds of the research findings, to use the data in an enrolment

management context in the identification of students "at risk" of academic exclusion. These students could then be targeted for remediation (the details of which are beyond the scope of this research) in an attempt to improve internal efficiency by means of decreasing rates of academic exclusion and repetition. Although it would have been desirable to statistically evaluate the effects of participation in the existing Academic Development Programme (ADP) and the Science Foundation Programme (SFP) on the final undergraduate outcome, this was not possible since ADP and SFP students were represented only within the lower faculty points groups.

All statistical analyses described in this study were performed using the StatSoft Statistica package. Each one of the four working databases was imported into Statistica, where numerical re-coding of variables to be used in statistical analyses was carried out prior to analysis; the text strings underlying each numeric variable were however displayed during the analyses and were displayed in the output from the programme. Details of the numerical coding employed are presented below in Exhibit 4.2 below:

Exhibit 4.2 : Numerical coding and recoding of data for statistical analysis

Database field	Contents	Numeric Coding Details
"SEX"	gender, alphabetic, M=male, F=female	M=1, F=0; in the binary coded dummy variable SEXN
"AUTHOR"	matriculation authority, alphabetic : ET, CO or OMA (see above)	Where matriculation authority was treated as a categorical variable, AUTHORN was recoded from AUTHOR in this way: DET=1, CO=2 and OMA =3. For discriminant analysis, 2 dummy variables (AUTDUM1 and AUTDUM2): DET = 0 1, CO = 1 0, OMA = 0 0 were set up.
"FACPOINT"	faculty points, numeric	no recoding necessary
"MATPTS"	Matric maths symbol, character	recoded as Swedish points per symbol
"ENGPTS"	Matric English symbol, character	recoded as Swedish points per symbol
"AGE"	age at June 1992 census, numeric	no recoding necessary
"DURATION"	duration of undergraduate enrolment at UCT, numeric	no recoding necessary
"LASTCODE"	last undergraduate progress code, character	"QUA" or "INQ" = 1, "REN" = 3, all others = 2

In certain of the analyses where relationships between categorical variables were explored (chi-square analyses and log-linear models), a new variable (FACGROUP) was created from groupings of the FACPOINT variable. Because the four faculties examined employ differential weightings in the calculation of faculty points and different selection policies in relation to faculty points, the range of the FACGROUP variable differed amongst the faculties. Details of the FACGROUP categories selected (which are consistent with those currently in use at UCT in the exploration of student performance in relation to matriculation criteria) are presented in Exhibit 4.3 below:

Exhibit 4.3 : Numeric recoding of faculty points data for categorical analysis

Faculty	"FACGROUP" Code :				
	1	2	3	4	5
Arts faculty points	<25	25-29	30-34	35-39	40+
Commerce faculty points	<45	45-49	50-54	55+	-
Science faculty points	<40	40-44	45-49	50-54	55+
Social Science faculty points	<25	25-29	30-34	35-39	40+

Unless otherwise stated, the "select cases" function within Statistica was employed to select those students with numeric LASTCODE = 1 or 3; in this way, those students who had left UCT voluntarily without completing their undergraduate degrees were excluded from the analysis. This step was necessary because of the wide range of possible reasons for voluntary withdrawal, and hence the academically diverse nature of the students involved. As will become clear in the discussion of voluntary withdrawals in Chapter 4, the last undergraduate progress codes amongst this group encompass a wide range of academic outcomes. Whilst voluntary withdrawal may have been linked to academic performance in some cases, many factors could have prompted the decision to withdraw amongst sound academic performers; in other words, it is unlikely that voluntary withdrawal could be predictable in terms of matriculation performance data above.

The analyses carried out within this study were therefore primarily focused on the statistical differentiation between successful students (or "qualifiers", with last undergraduate progress code = "QUA" or "INQ") and those who had been refused readmission to UCT because of poor academic performance (with a last undergraduate progress code of "REN"), and the characterisation and possible prediction of members of these two groups. In all statistical tests and models, cases with missing data were excluded from the analyses. The alpha level of significance was set to 0.05 for all tests; statistical significance was indicated with a "YES" ($\alpha \leq 0.05$) or a "NO" ($\alpha > 0.05$).

(i) Chi-square analysis and log-linear modelling

Chi-square analysis was carried out in order to explore possible two-way relationships between the last undergraduate progress code (LASTCODE) and a range of categorical variables contained within each of the 4 working databases. Chi-square analysis was used in particular to detect significant differences in the frequencies of the variable LASTCODE as cross-tabulated against the variables AUTHORN (the numerically recoded matriculation authority group) and FACGROUP (faculty points group, see above) within each of the 4 faculties.

Where significant associations between particular independent (design) variables and the dependent (response) variable LASTCODE were detected, a log-linear analysis was carried out in order to formulate a model describing the relationships between the frequencies in the multi-way observed frequency table.

Within Statistica's Log-Linear module, the relationship between the variables in the analysis was initially probed by exploring the significance of all k-factor interactions (where k equals the number of variables in the analysis). The results of this evaluation, which tests the null hypotheses that (i) there are no significant main effects, (ii) that there are no significant two-way interactions and (iii) that there are no significant three-way interactions between the variables, were used to pre-determine the likely interaction levels within the model. Significant p-values obtained for k=2 (i.e. two-way interactions) or k=3 (i.e. three way interactions) indicated that these higher level interactions significantly improved the fit of the log-linear model.

Once the probable level of interactions between the variables had been determined, and where particular two-way interactions were found to be significant, an examination of the marginal and partial associations[20] amongst the variables involved was carried out in order to identify probable interactions within the model. Based on the evaluations of all k-factor interactions and the tests of marginal and partial association, a "manual" best fit model was identified.

Thereafter, the log-linear module's automatic model fitting option, which determines the best fitting model which would describe the data in the observed multi-way table[21], was applied in order to validate the manually fitted model. An iterative algorithm is employed in automatic

[20] The test of partial associations evaluates the effects of removing a particular interaction from the saturated model (i.e. a model incorporating all interactions at the level determined by fitting all K-factor interactions); the test of marginal associations tests the significance of adding individual interactions to a model without any interactions.

[21] Within Statistica's Log-Linear module, the constant *delta* (which is set at 0.5 by default) is added to all frequencies in the observed table prior to analysis; this correction is recommended when the multi-way table contains a number of cells with frequencies below 10.

model fitting; this algorithm initially fits a model including the main effects of individual variables (with no interactions) only; if the main effects model does not fit (i.e. chi-square analysis indicates that the expected cell frequencies under the model differ significantly from the observed cell frequencies), the program will fit and test a model including all two-way interactions. The process would be repeated fitting all three-way interactions if the model including all two-way interactions did not fit the observed data. Once a model which fits the data is obtained, the program tests all interactions in the model, and ultimately includes only those necessary to fit the data in the observed multi-way table. The automatic and "manual" best-fit models were compared prior to further evaluation of the best-fit model.

Once a best fit log-linear model had been obtained for a particular faculty cohort, the goodness of fit of the model was tested using chi-square analysis; in this analysis an insignificant chi-square value (i.e. $p>.05$) indicates that there is no significant difference between expected cell frequencies under the model and the observed frequencies in the multi-way table and thus that the model provides an adequate description of the data within the observed multi-way table. The best-fit model was further evaluated by means of an inspection of the residual (i.e. observed minus expected) frequencies and by means of plotting observed frequencies against fitted frequencies under the best-fit model. Where interactions between the design variables AUTHORN and FACGROUP were included in the best fit model, the effects of omitting such interactions from the model were tested hierarchically by examining the goodness of fit of a model excluding the interaction. Similarly, where the interaction between a design variable and the response variable was not one of the effects in the best-fit model, the difference (based on the statistical significance of the differential chi-square statistics) between the fit of a model plus this effect and the best-fit model was assessed. If the inclusion of the missing interaction resulted in a statistically significant improvement in the model fit, then this effect would be included in the model.

Finally, the relationships between the design variables and the response variable (LASTCODE) detected by the best fit model were interpreted via visual inspection of the marginal tables representing the significant interactions in the model. Details of the development of the log-linear model for each faculty, the *post hoc* evaluation of the best fit model and the composition of the marginal tables are included within Appendix 1.

(ii) Multiple discriminant analysis

Multiple discriminant analysis was applied in order to detect significant discriminator variables of the grouping variable LASTCODE amongst a range of potential (ratio-scaled) matriculation-related predictor variables contained within the working database for each faculty. A secondary aim of this type of analysis was to generate, for each faculty cohort, a classification function which could be used to predict the group membership of new cases

(such as prospective new students) or of new cohorts of FU students embarking on 3-year bachelors' degrees within each of the 4 faculties examined. In order to include the categorical gender and matriculation authority variables in the analysis, dummy variables were set up in accordance with the rule of using *k - 1* dummy variables, where *k* is the number of levels in the categorical variable (Anderson, Sweeney and Williams, 1993:577). Accordingly, the two-level gender variable was replaced with a binary coded numeric variable (SEXN) and two dummy variables (AUTDUM1 and AUTDUM2) were set up to represent the three-level matriculation authority groupings within the variable AUTHOR - see Exhibit 4.2 above.

In preparation for multiple discriminant analysis, the potential predictor variables within each dataset were screened for inter-correlation. Where strong correlations between potential predictor variables were detected (e.g. between FACPOINT and MATPTS) the more important variable (FACPOINT in this case, which was fundamental to the analysis) was included in the analysis, unless the significance or predictive power of the resultant model was compromised in so doing. The group means (i.e. the means for predictor variables for each group within the dependent variable LASTCODE) of the potential predictor of all variables to be included in the analysis, as well as box-and-whisker plots by group within the grouping variable LASTCODE, were then inspected in order to assess which of these variables could potentially be included in the multiple discriminant model resulting from the analysis. Details of these evaluations for each of the 4 faculty cohorts examined are presented in Appendix 2.

Within Statistica's Discriminant Analysis module, a forward step-wise analysis was then performed for each faculty, using the variable LASTCODE as the grouping variable. F to enter for the analysis was set to 3, and Wilks' *lambda* controlled the entry order. Discriminant function analysis summaries of the models generated for each of the 4 faculties, and tests of the significance of these models, are shown in Appendix 2. These summaries indicate (a) whether or not there are significant differences between the "qualified" ("QUA") and "excluded" ("REN") groups within the LASTCODE variable and, if so, (b) the relative importance of each of the predictor variables in discrimination between groups.

The classification matrices pertaining to each multiple discriminant model were then viewed in order to assess the predictive power of the models in terms of predicting group membership. These classification matrices are presented in Appendix 2, and are discussed within the Results chapter (Chapter 5) along with the probable reasons for the relatively poor classification capabilities of the multiple discriminant models. Initial inspection of the data within all four cohorts revealed that there were considerably more cases in the "QUA" group; it was therefore specified in the analysis that the *a priori* probabilities should therefore be proportional to the size of each group. It is important to recognise that basing the predictive power of the classification functions on the same data which was used to set up the

discriminant functions can result in an overestimation of the predictive power of the model, which is thus based on *post hoc* predictions. It is therefore important to test promising classification functions on new cases (with known group membership) so as to more accurately assess the predictive capabilities of these expressions. This is commonly achieved by deriving the classification functions from half of the cases in the dataset, and subsequently testing the predictive capacity of these functions on the balance of the cases in the dataset. This was not possible within the current analyses because the substantial differences between the proportions of "QUA" and "REN" group members (and hence the relatively small number of "REN" group members), and the large number of instances where missing data resulted in the casewise deletion of cases, meant that there were too few "REN" cases to permit the division of the dataset.

The results of both the descriptive investigations and the statistical analyses for each of the cohorts are presented in Chapter 5.

CHAPTER 5

DESCRIPTIVE AND STATISTICAL ANALYSIS OF ACADEMIC PERFORMANCE OF SELECTED COHORTS

This chapter presents a summary of the results of the descriptive and the statistical analyses carried out for each of the four selected cohorts; the results are presented as a separate analysis of each cohort. The analysis of each cohort consists of a descriptive analysis of the graduation rates, the time to degree, the rates of academic exclusion and the voluntary drop-out rates in relation to the matriculation (matric) authority and faculty points groups outlined in Chapter 4, followed by an examination of the specific results of the log-linear and multiple discriminant analyses for each cohort. An inter-faculty summary of the key findings of concludes this chapter.

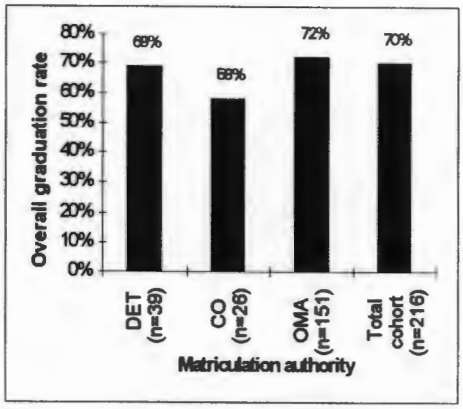
5.1 ANALYSIS OF THE BA COHORT, FACULTY OF ARTS

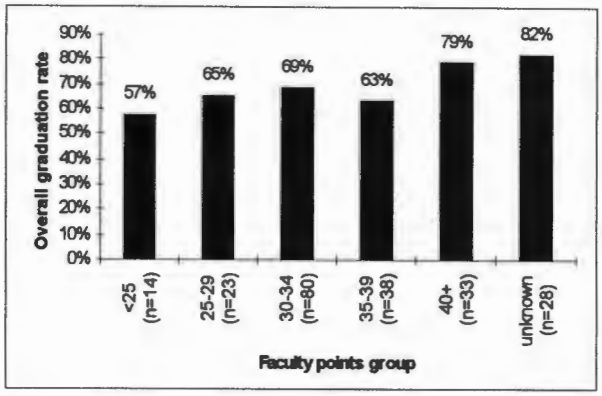
5.1a Descriptive analysis

Of the 258 (first-time entering) FU students entering the BA program in the Faculty of Arts in 1992, a considerable proportion (42 students, or 16% of the cohort) transferred to other faculties at UCT; the majority of these students (81%) left the Arts faculty after one year of study.

The graduation rate amongst the remaining 216 students (referred to from now on as the BA cohort) as a function of years of enrolment, matriculation authority and faculty points group is presented in Exhibit 5.1 below. A comparison of the graduation rates as a function of matric authority shows that there was little difference in the ultimate undergraduate success rates of the ex-DET and Transkei students (i.e. the "DET" group - 69%) and those who wrote the examinations of the previously white matriculation authorities (i.e. the "OMA" group - 72%) students; the graduation rate of 58% amongst those who had written the Coloured Senior Certificate (i.e. the "CO" group) was however considerably lower than within the other two groups. The graduation rate tended to increase with increasing faculty points groups, from 57% amongst the lowest group to 79% amongst those who had entered the Faculty with 40 or more faculty points. The increase in graduation rates as a function of increasing faculty points groups was, however, not uniform with a particularly low rate of graduation (63%) occurring amongst the 35-39 faculty points group:

Exhibit 5.1 : Comparison of graduation rates in relation to matriculation authority and faculty points group (BA cohort)





Whilst overall graduation rates within the Faculty were of a high level, internal efficiency in the production of graduates was reduced by the large proportions of students who took 4 years or more to complete their BA degrees. Exhibit 5.2, which summarises the time to degree as a function to matric authority and faculty points group, shows that only 53% of the Arts cohort had graduated within the minimum 3 year period; particularly small proportions of the DET students (41%) and CO students (38%) had graduated within 3 years. Similarly, small proportions of those entering the faculty with less than 30 faculty points (36% of the <25 points group and 26% of the 25-29 points group) were found to have completed their degrees within 3 years.

Exhibit 5.2 : Time to degree in relation to matriculation authority and faculty points group (BA cohort)

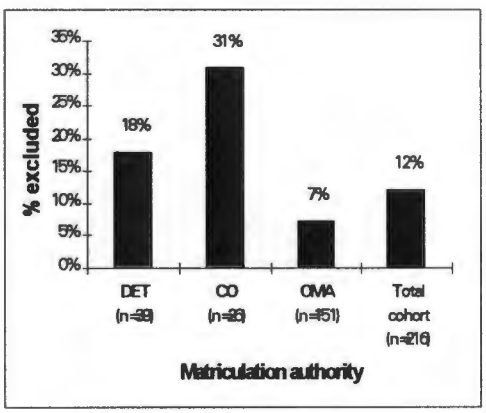
	MATRIC AUTHORITY :			TOTAL COHORT	FACULTY POINTS GROUP :					
	DET	CO	OMA		<25	25-29	30-34	35-39	40+	unknown
No. entering Faculty	39	26	151	216	14	23	80	38	33	28
No. graduating after:										
3 years	16	10	89	115	5	6	44	20	22	18
	41%	*38%*	*59%*	*53%*	*36%*	*26%*	*55%*	*53%*	*67%*	*64%*
4 years	9	5	18	32	3	7	11	3	3	5
	23%	*19%*	*12%*	*15%*	*21%*	*30%*	*14%*	*8%*	*9%*	*18%*
5 years	2	0	2	4	0	2	0	1	1	0
	5%	*0%*	*1%*	*2%*	*0%*	*9%*	*0%*	*3%*	*3%*	*0%*
Overall graduation rate	27	15	109	151	8	15	55	24	26	23
	69%	*58%*	*72%*	*70%*	*57%*	*65%*	*69%*	*63%*	*79%*	*82%*

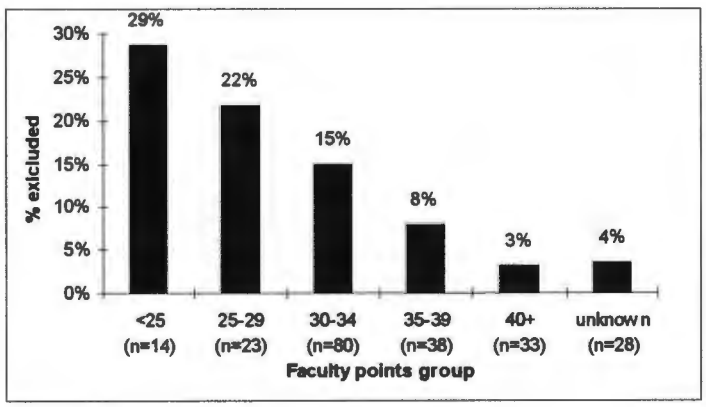
Chi-square analysis revealed that there was a marginally significant association between the time to degree (categorised into two groups, *viz.* three years or more than three years) and the matric authority group amongst the Arts cohort graduates (Pearson chi-square = 6.644 at df=2, with p=.03609); a considerably larger proportions of the OMA group (59%, in comparison with 41% of the DET group and 38% of the CO group) had completed their degrees within the minimum three-year period. There is, however, a highly significant

association between the faculty points group and the time to degree (in terms of the three years or more than three years categorisation) amongst the graduates within this cohort (Pearson chi-square = 13.32397 at df=4, with p=.00980), with larger proportions of the graduates within the higher faculty point groupings completing their degrees within three years. Conversely, graduates within the lower faculty point groupings tended to take longer than those within the higher faculty points groupings to complete their degrees: 21% of the <25 faculty points group and 39% of the 25-29 faculty points group were found to have taken more than three years to complete their basic degrees.

Rates of student attrition due to academic exclusion are depicted in Exhibit 5.3 below. There was a low overall rate of academic exclusion amongst the cohort (12%), but the rate of exclusion was seen to vary as a function of both matric authority and faculty points group. The rates of exclusion amongst DET and CO students (18% and 31% respectively) were considerably higher than amongst OMA student (7%). The rate of academic exclusion was also seen to decline progressively from a level of 29% amongst the <25 faculty points group to only 3% amongst those with 40 or more faculty points.

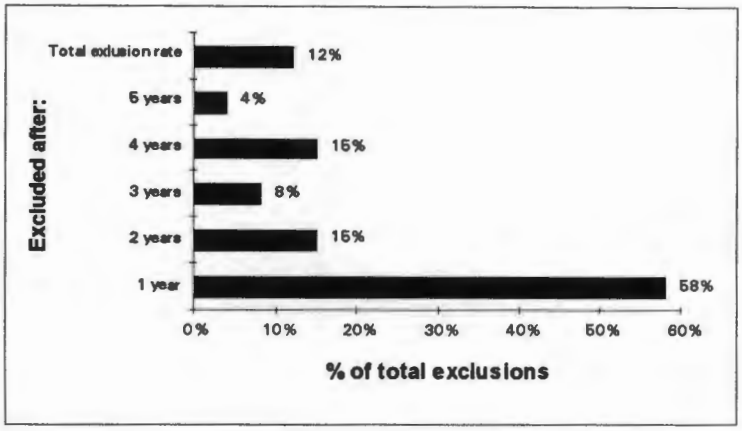
Exhibit 5.3 : Academic exclusion rates in relation to matriculation authority and faculty points group (BA cohort)





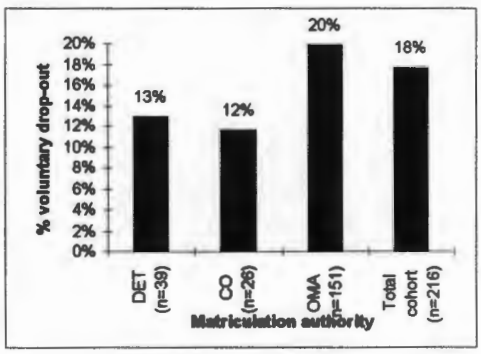
The largest proportion of the academic exclusions within the cohort (58%) occurred at the end of the first year of enrolment (see Exhibit 5.4). Consequently, 42% of those ultimately excluded on academic grounds were enrolled for 2 full years or more prior to exclusion.

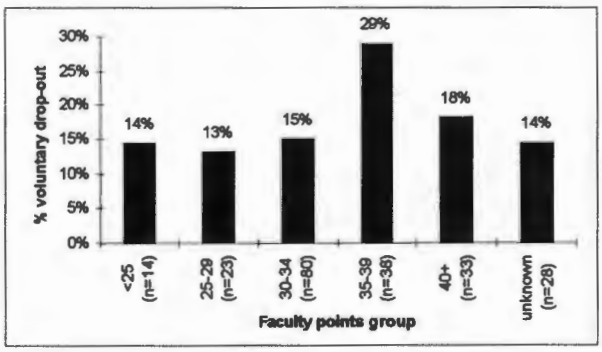
Exhibit 5.4 : Academic exclusions in relation to period of enrolment (BA cohort)



The other component of student attrition is that resulting from voluntary drop-out. Exhibit 5.5 indicates that the overall drop-out rate within this cohort (18%) was considerably higher than the academic exclusion rate. The pattern observed amongst drop-out rates in relation to matric authority and faculty points groups also differed markedly from that seen in Exhibit 5.4: the drop-out rate was particularly high amongst OMA students (20% of the OMA cohort), and within the higher 35-39 and 40+ faculty points groups (29% and 18% respectively).

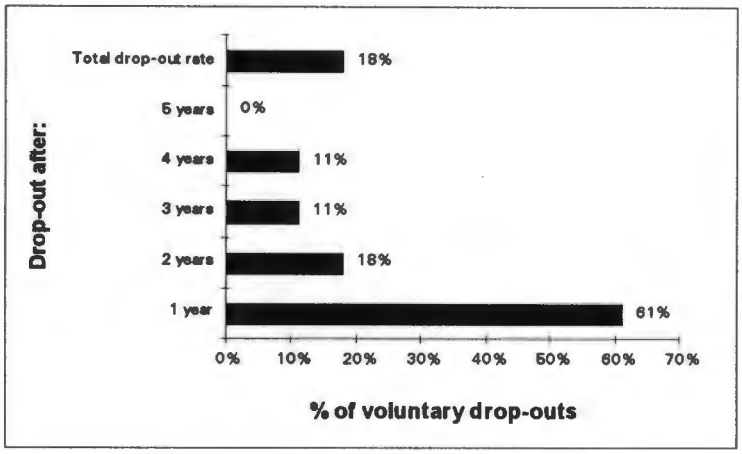
Exhibit 5.5 : Voluntary drop-out rates in relation to matriculation authority and faculty points group (BA cohort)





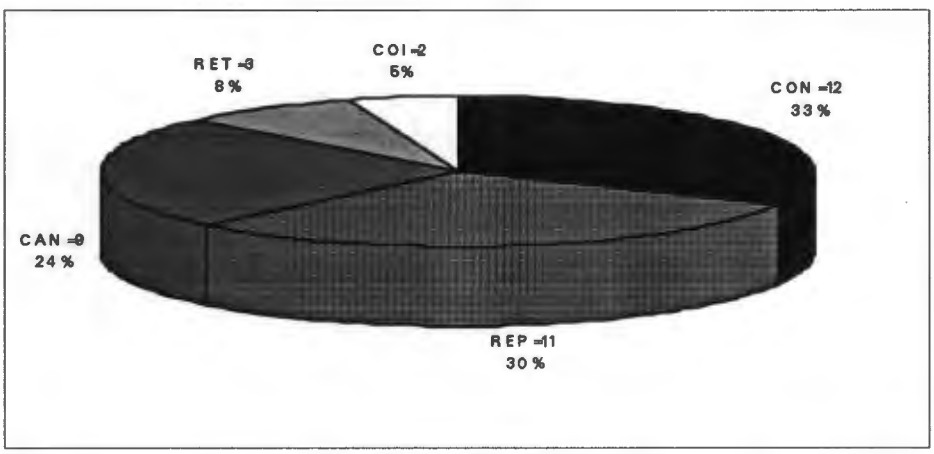
Voluntary drop-out was seen to occur most commonly at the end of the first year of study (Exhibit 5.6), but was also common after 2 or more years of study: 39% of the voluntary drop-outs were found to have occurred following two years of study.

Exhibit 5.6 : Voluntary drop-out in relation to duration of enrolment (BA cohort)



A summary of the last academic progress codes amongst the group of voluntary drop-outs is shown in Exhibit 5.7. This summary indicates that academic difficulty may have been a factor in 43% of the voluntary drop-outs (those whose last undergraduate records reflected a "REP", "RET" or "COI" code), but that other factors (such as financial difficulties or social integration problems) may have led the remainder of the drop-outs to either withdraw from UCT (those with progress code "CAN") or to fail to re-register at UCT in order to complete their degrees (i.e. those with a final progress code of "CON").

Exhibit 5.7 : Last undergraduate promotion codes amongst voluntary drop-outs (BA cohort)



Only one member of the BA cohort was found to be continuing with his/her undergraduate studies within the Faculty of Arts in 1997.

5.1b Statistical analysis

i) Log-linear modelling

Log-linear modelling was used to detect interactions within a three way cross-tabulation between the design variables FACGROUP (faculty points group) and AUTHORN (matric authority group) and the "QUA" or "REN" groups response variable LASTCODE within the BA cohort. The development of the log-linear model for this cohort is described in detail in Appendix 1 of this report, and the numeric coding of the variables in the analysis is explained in Chapter 4. The analysis returned a best-fit log-linear model which incorporated (i) the interaction between the design variables FACGROUP and AUTHORN and (ii) the interaction between the design variable AUTHORN and the response variable LASTCODE. The interaction between the design variables resulted from the fact that the faculty points groups were not evenly represented amongst the matriculation authority groups within the cohort; students within the lower faculty points groups were largely DET students. Further analysis of the AUTHORN*LASTCODE interaction (which indicated that there was a significant interaction between the rates of graduation and academic exclusion varied in relation to matriculation authority) revealed that the graduation:academic exclusion ratio was considerably higher amongst the OMA group (7.24) than amongst either the DET group (3.0) or the CO group (1.7). The graduation:academic exclusion ratio amongst the cohort as a whole (4.2) was clearly forced downwards by the relatively low rates of graduation and high rates of academic exclusion within the DET and CO matric authority groups.

Log-linear analysis revealed that there was no significant interaction between the categorical variable FACGROUP and the response variable LASTCODE. In other words, the relative rates of graduation and of academic exclusion were not significantly influenced by the FACGROUP variable.

The results of the log-linear analysis also indicated that there was no significant 3-way interaction (i.e. FACGROUP*AUTHORN*LASTCODE) within the 1992 Arts cohort; this implied that the AUTHORN*LASTCODE interaction was not modified by the FACGROUP (faculty points group) membership of students within this cohort. In other words, the relatively high rates of academic exclusion amongst the DET and particularly the CO students were not significantly affected by the faculty points group membership of these students.

ii) Multiple discriminant analysis

Multiple discriminant analysis was carried out firstly to detect predictors of "QUA" or "REN" group membership within the variable LASTCODE from a range of independent variables including SEXN (gender), AUTDUM1 (where the CO matric authority was coded as 1), AUTDUM2 (where the DET matric authority was coded as 1), FACPOINT (the ungrouped faculty points for each student), ENGPTS (the Swedish points equivalent of the matric

English symbol) and MATPTS (the Swedish points equivalent of the matric mathematics symbol); the numeric coding of the variables in the analysis is described in Chapter 4. A second aim of the multiple discriminant analysis was to derive classification functions which could be used to classify new cases (i.e. new students or applicants) as either "QUA" or "REN" group members based on the significant discriminator variables; such classification functions would be potentially useful in informing faculty admissions policies and in targeting students at risk of academic exclusion for remedial intervention. Prior inspection of the inter-correlations between the potential predictor variables indicated that there was a moderate to strong correlation between the variables FACPOINT and MATPTS (R=0.73); because the variable FACPOINT was a key variable in this analysis, the correlated variable MATPTS (which is the Swedish points equivalent of the matric Mathematics symbol) was subsequently excluded from the multiple discriminant analysis. The development of the multiple discriminant model is detailed in Appendix 2.

The multiple discriminant analysis summary indicated that there was a significant discrimination (F=8.6319 with p<.0003) between "QUA" and "REN" groups within the ultimate undergraduate academic outcome variable LASTCODE, and that the variables FACPOINT (the ungrouped faculty points score for each student) and AUTDUM1 (i.e. the CO matriculation authority) were the best predictors of "QUA" or "REN" group membership. However, the high Wilks' *lambda* for the model (0.89554) suggested that the model possessed poor discriminatory power with regard to "QUA" or "REN" group membership.

Evaluation of the classification functions associated with the multiple discriminant model confirmed this. Examination of the classification matrix showed that although application of the classification functions resulted in the correct classification (in terms of their "QUA" or "REN" group membership) of 84% of the cases in the analysis, the accuracy of classification was particularly poor amongst the true members of the "REN" group; 99% of the "QUA" group and only 4% of the considerably smaller "REN" group were found to have been classified correctly. In other words, 96% of the members of the Arts cohort who had been excluded from UCT on academic grounds, were incorrectly classified as graduates in terms of the classification functions. Bearing in mind that the classification discussed here is a *post hoc* classification (i.e. a re-classification of the cases from which the classification functions were derived) and that application of the classification functions to new cases would result in an even poorer predictive accuracy, it appears that the multiple discriminant model has little potential in the predictive classification in terms of detecting potential "REN" group membership amongst prospective BA students.

5.1c BA cohort - Summary

The descriptive analysis of the 1992 BA cohort revealed that 70% of the cohort had successfully completed a basic degree within the Faculty of Arts by the end of the 1996

academic year. There were, however, notable differences in the graduation rates amongst the different matriculation authorities represented within the cohort: a particularly small proportion of the CO matriculants (58%), in comparison with 69% of the DET students and 72% of the OMA students had graduated within the faculty by the end of 1996. The proportion of graduates tended to increase amongst increasing faculty points groups (from 57% within the <25 points group to 79% of the 40+ faculty points group), although there was a particularly low graduation rate (63%) amongst the 35-39 faculty points group. It appeared that a particularly low proportion of the CO students (38%, in comparison with 41% of the DET students and 59% of the OMA students) had graduated within three years of enrolling at UCT, and chi-square analysis showed that there was a marginally significant association between the time to degree and the matriculation authority group membership of students within the BA cohort. Chi-square analysis however revealed a highly significant relationship between faculty points group membership and the time to degree; the proportion of students who completed their degrees within three years increased from 36% within the <25 faculty points group to 67% within the 40+ faculty points group.

The rate of academic exclusion within this cohort (12% overall) was found to be especially high amongst the CO students (31%) and particularly low amongst the OMA students (7%). The rate of academic exclusion was also seen to decline from 29% amongst the <25 faculty points group to only 3% of amongst the 40+ faculty points group. Conversely, the rate of voluntary drop-out (which was 18% for the whole cohort) appeared to be considerably higher within the OMA group (20%) than within either the DET or the CO groups. Voluntary drop-out was also more common amongst the higher faculty points groups: 29% of the 35-39 faculty points group and 18% of the 40+ faculty points groups had dropped out of UCT in good academic standing. The majority of the both the voluntary drop-outs (61%) and the academic exclusions within this cohort (58%) were found to have occurred following one year of enrolment at UCT.

Log-linear modelling of the three-way relationship between the matriculation authority, faculty points group and final undergraduate outcome variable revealed that there was a significant association between the matriculation authority and the final undergraduate outcome, and that the relatively low ratios of graduation:academic exclusion ratios amongst the DET and CO matriculation authority groups had tended to diminish the graduation:academic exclusion ratio within the cohort as a whole.

The results of a forward, stepwise multiple discriminant analysis carried out in order to detect significant predictors of the final undergraduate academic outcome amongst a range of independent variables revealed that the ungrouped faculty points variable and membership of the CO matriculation authority group were the most significant predictor variables. The high Wilks' *lambda* for the multiple discriminant model, and the relatively poor predictive accuracy of the classification functions with regard to membership of the group of academic

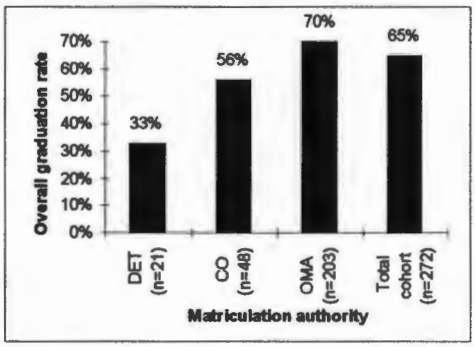
exclusions however indicated that the multiple discriminant analysis would be of little use in predicting the likely group membership of new data cases (e.g. potential students, or those within other year of entry cohorts).

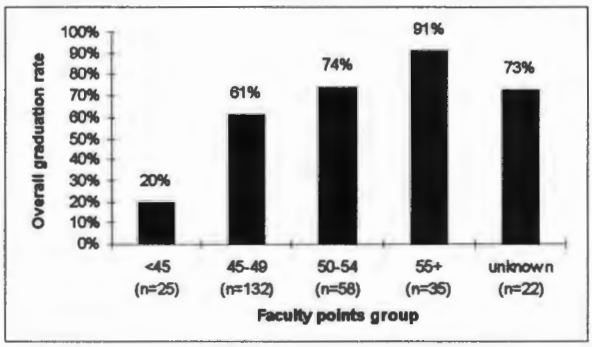
5.2 ANALYSIS OF THE BCom COHORT, FACULTY OF COMMERCE

5.2a Descriptive analysis

A total of 285 FU students entered the BCom program in 1992; of these only 13 (5% of the original cohort) transferred to other faculties at UCT. Exhibit 5.8, which depicts the graduation rate amongst the remaining 272 students (henceforward referred to as the FU BCom cohort) as a function of matric authority and of faculty points group, indicates that there was an overall graduation rate of 65% amongst this cohort. The graduation rate amongst DET students (33%) was however considerably lower than the rate for the whole cohort (65%), whilst that amongst OMA students (70%) was substantially higher than this figure. The graduation rate generally increased with increasing faculty points group from a level of 20% amongst the <45 points group to a high of 91% amongst the 55+ faculty points group. Graduation rates within the cohort therefore varied dramatically in terms of both the matric authority and the faculty points groupings within this cohort.

Exhibit 5.8 : Comparison of graduation rates in relation to matriculation authority and faculty points group (BCom cohort)





The analysis of time to degree as a function of matric authority and faculty points grouping presented in Exhibit 5.9 indicates that only 36% of the cohort graduated within the 3-year minimum enrolment period and that 29% of the cohort therefore graduated after four or more years of enrolment within the Commerce Faculty. Whilst 43% of the OMA students and 21% of the CO students had graduated following three years of enrolment, the minimum time to degree amongst the DET students was found to be four years. Examination of the time to degree as a function of faculty points group revealed that whilst 69% of the 55+ group had completed their degrees after three years, only 1 (4%) of the <45 points group and 29% of the 45-49 points group had graduated within the minimum enrolment period. This analysis indicated that although there was a relatively low attrition rate amongst the 1992 FU BCom

cohort, efficiency within this degree programme was compromised by the evidently low rate of graduation within the minimum 3 year enrolment period.

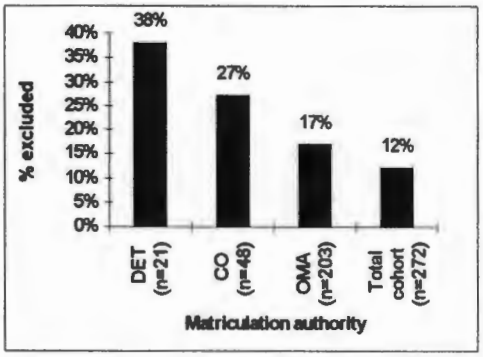
Exhibit 5.9 : Time to degree in relation to matriculation authority and faculty points group (BCom cohort)

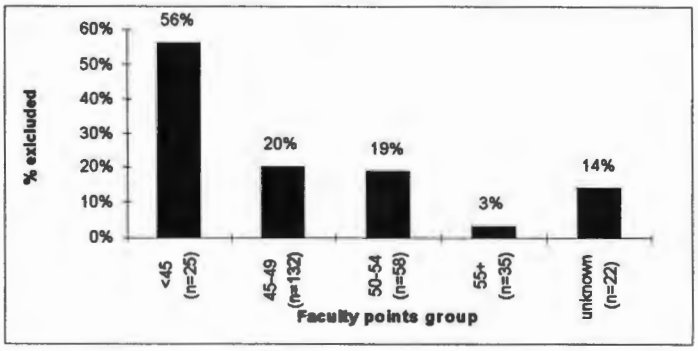
	MATRIC AUTHORITY :			TOTAL COHORT	FACULTY POINTS GROUP :				
	DET	CO	OMA		<45	45-49	50-54	55+	unknown
No. entering Faculty	21	48	203	272	25	132	58	35	22
No. graduating after:									
3 years	0	10	87	97	1	38	27	24	7
	0%	*21%*	*43%*	*36%*	*4%*	*29%*	*47%*	*69%*	*32%*
4 years	3	15	45	63	2	36	12	6	7
	14%	*31%*	*22%*	*23%*	*8%*	*27%*	*21%*	*17%*	*32%*
5 years	4	2	10	16	2	6	1	2	2
	19%	*4%*	*5%*	*6%*	*8%*	*5%*	*2%*	*6%*	*9%*
Overall graduation rate	7	27	142	176	5	80	40	32	16
	33%	*56%*	*70%*	*65%*	*20%*	*61%*	*69%*	*91%*	*73%*

Chi-square analysis revealed that there was a highly significant relationship between the matric authority grouping and the time to degree (categorised into two groups, *viz.* three years or more than three years) amongst the graduates within the BCom cohort (Pearson chi-square = 14.335 at df=2, with p=.00077). Similarly, chi-square analysis showed that there was a significant association between the faculty points grouping and the time to degree (in terms of the three year and more than three year categories) amongst the graduates within this cohort (Pearson chi-square = 10.478 at df=3, with p=.01492). These analyses indicate that both the faculty points group and the matric authority group contributed to the time to degree amongst the graduates within this cohort. Graduates who entered the Faculty within the lower faculty points groups clearly took longer than those in the higher faculty points groups to complete their degrees. At the same time, considerably larger proportions of the OMA graduates (in comparison with those from the DET and CO matric authority groups) completed their BCom degrees within the minimum three-year period.

Exhibit 5.10 shows the overall rate of attrition due to academic exclusion within this cohort (21%), and compares exclusion rates in relation to matric authority and faculty points groups. The rate of academic exclusion was clearly highest (38%, which is greater than the graduation rate for this group) amongst DET students, and lowest (17%) amongst OMA students. Likewise, the rate of academic exclusion declined from 56% amongst the <45 faculty points group to a level of only 3% amongst the 55+ group.

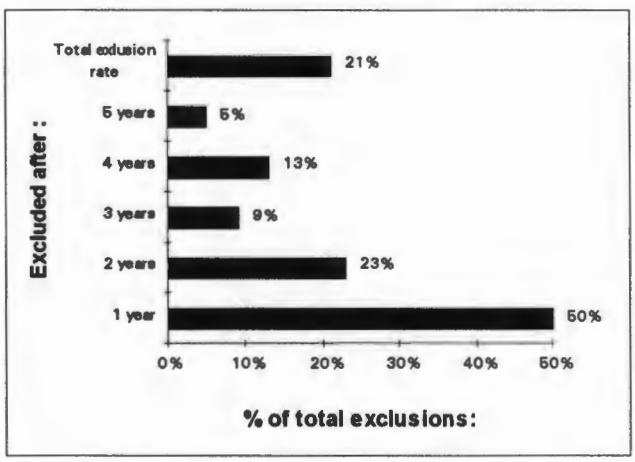
Exhibit 5.10 : Academic exclusion rates in relation to matriculation authority and faculty points group (BCom cohort)





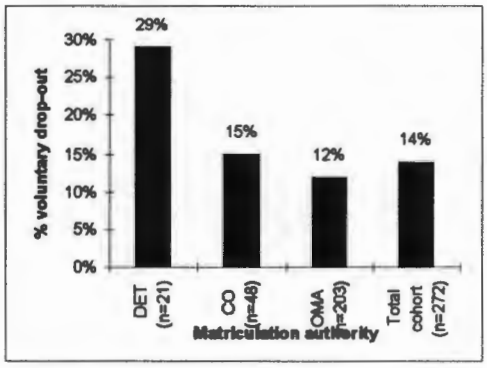
The relatively high rate of academic exclusion within the cohort and its uneven distribution are clearly cause for concern; moreover, it appears that half of these exclusions occurred following 2 or more years of enrolment within the Commerce faculty (Exhibit 5.11).

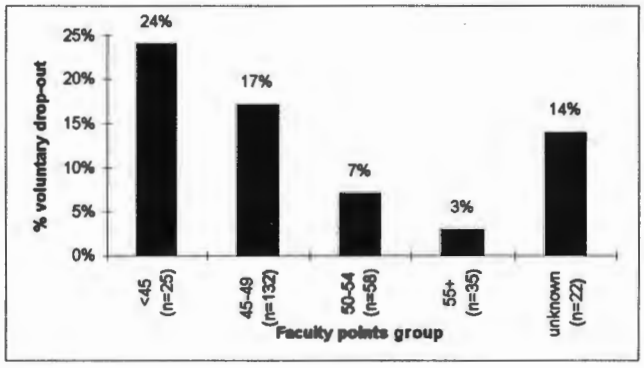
Exhibit 5.11 : Academic exclusions in relation to duration of enrolment (BCom cohort)



The voluntary drop-out rate amongst the BCom cohort (as a function of matric authority and faculty points group) is shown in Exhibit 5.12. The drop-out rate (14% for the whole cohort) varied both in relation to matric authority and faculty points group: the highest rate of voluntary drop-out (29%) occurred amongst DET students whilst drop-out was most common (24%) amongst the <45 faculty points group.

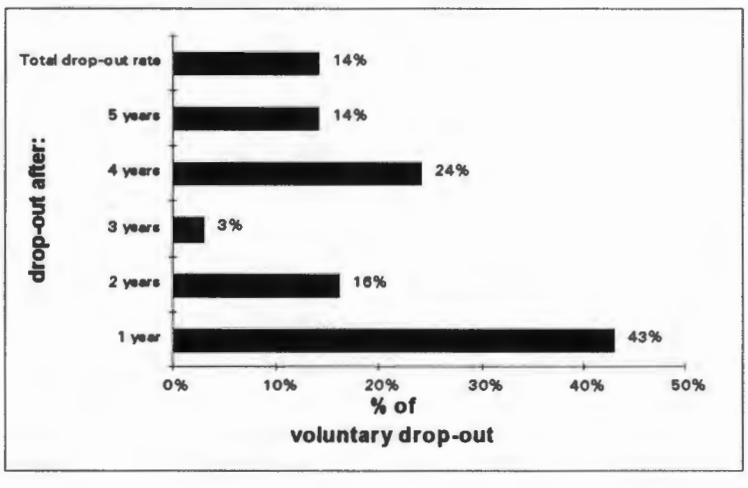
Exhibit 5.12 : Rates of voluntary exclusion in relation to matriculation authority and faculty points group (BCom cohort)





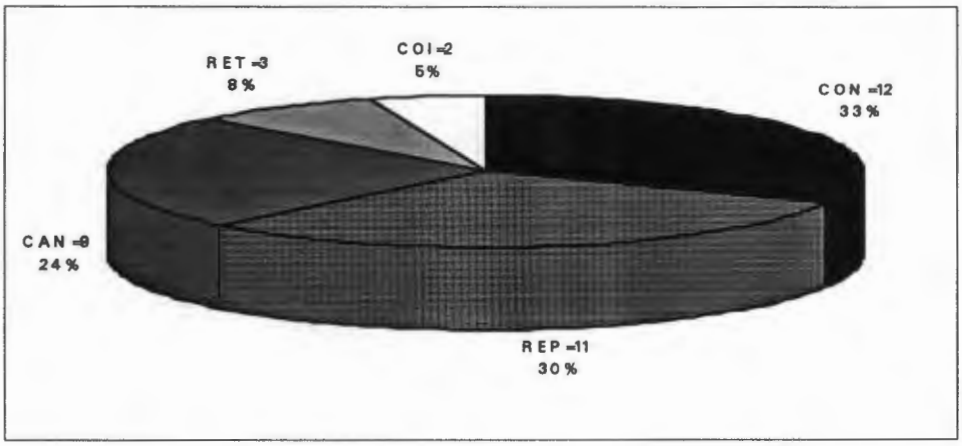
Although the largest proportion (43%) of these drop-outs occurred at the end of the first year of enrolment, a substantial proportion (16%) occurred after two years of enrolment and a large proportion (24%) of the voluntary drop-outs took place after four years of enrolment in the Commerce faculty (see Exhibit 5.13 below).

Exhibit 5.13 : Voluntary drop-out in relation to duration of enrolment (BCom cohort)



The last undergraduate progress codes achieved by the voluntary drop-outs are summarised in Exhibit 5.14. This analysis indicates that voluntary drop-out was possibly related to academic difficulties in at least 40% of such cases (i.e. those failing to re-register having achieved a "REP", "RET" or "COI" code); however, the quantitative analysis employed here does little to explain the failure to re-register amongst the 32% of drop-outs whose last undergraduate records reflected the code "CON", or amongst the 24% of the voluntary drop-outs leaving UCT after cancelling their registration (with a last undergraduate progress code of "CAN").

Exhibit 5.14 : Last undergraduate progress codes amongst voluntary drop-outs (BCom cohort)



Only 3 members (1.1%) of the 1992 BCom cohort were found to be enrolled as undergraduates within the Commerce Faculty in 1997.

5.2b Statistical analysis

i) Log-linear modelling

The development of a log-linear model, which aimed to detect statistically significant interactions between the variables AUTHORN (matric authority group), FACGROUP (faculty points group) and "QUA" or "REN" group membership within the response variable LASTCODE, within a three-way cross-tabulation representing the BCom cohort, is described in detail in Appendix 2. The numeric coding pertaining to the variables in the analysis is explained in Chapter 4. The analysis returned a best-fit log-linear model which incorporated (i) the interaction between the design variables FACGROUP and AUTHORN and (ii) the interaction between the design variable FACGROUP and the response variable LASTCODE. The interaction between the design variables resulted from the fact that the faculty points groups were not evenly represented amongst the matriculation authority points groups within the cohort. Further analysis of the marginal tables pertaining to the FACGROUP*LASTCODE interaction (which indicated that the rates of graduation and academic exclusion varied in relation to faculty points group) revealed that the graduation:academic exclusion ratio varied from 0.42 for FACGROUP=1 (<45 points), to 2.86 for FACGROUP=2 (45-49 points), to 3.56 for FACGROUP=3 (50-54 points) to 13.4 for FACGROUP=4 (55 points and above). The

graduation:academic exclusion ratio was less than 1 within the <45 faculty points group only, i.e. the rate of academic exclusion within this group exceeded the graduation rate. The overall graduation:academic exclusion ratio for the cohort as a whole (2.8) was clearly forced downwards by the relatively low ratios for the large (n=148) 45-49 faculty points group and the <45 faculty points group.

Chi-square analysis showed that there was a significant two-way association between the variables AUTHORN and LASTCODE (chi-square=10.35 with df=2 and p=.00566), and a highly significant association between the variables FACGROUP and LASTCODE (chi-square=33.236 with df=3 and p=.0000). Log-linear analysis however revealed that there was no significant interaction between the categorical variable AUTHORN and the response variable LASTCODE within the three-way AUTHORN*FACGROUP*LASTCODE table.

The results of the log-linear analysis also indicated that there was no significant 3-way (i.e. FACGROUP*AUTHORN*LASTCODE) interaction within the 1992 BCom cohort; this implied that the FACGROUP*LASTCODE interaction was not modified by the AUTHORN (matric authority group) membership of students within this cohort. In other words, the relatively low rates of graduation amongst the <45 faculty points group the 45-49 faculty points group were not affected by the matric authority group membership of these students.

ii) Multiple discriminant analysis

The development of the multiple discriminant model for the 1992 BCom cohort is detailed in Section 6b of Appendix 2. Multiple discriminant analysis was carried out in order to select predictors of "QUA" or "REN" group membership within the variable LASTCODE from a list of independent variables. A second aim of the multiple discriminant analysis was to derive classification functions which could be used to classify new cases (i.e. new students or potential students) as either "QUA" or "REN" group members. Prior inspection of the inter-correlations between the potential predictor variables - SEXN (gender), AUTDUM1 (where the CO matric authority was coded as 1), AUTDUM2 (where the DET matric authority was coded as 1), FACPOINT (the ungrouped faculty points for each student), ENGPTS (the Swedish points equivalent of the matric English symbol) and MATPTS (the Swedish points equivalent of the matric mathematics symbol) - indicated that there was a moderate correlation between the variables FACPOINT and MATPTS. However, experimental exclusion of the less important MATPTS variable was found not to affect either Wilks' lambda for the model, or the entry of variables into the model. Both FACPOINT and MATPTS were therefore included in the analysis. The numeric coding of the variables in the analysis is explained in Chapter 4.

The multiple discriminant analysis summary indicated that there was a significant discrimination (F=35.9222 with p<.000) between "QUA" and "REN" groups within the ultimate

undergraduate academic outcome variable LASTCODE, and that the variable FACPOINT alone (the ungrouped faculty points score for each student) was a significant discriminator of "QUA" or "REN" group membership. However, the relatively high value observed for Wilks' *lambda* for the model (0.859393) suggested that the model possessed weak discriminatory power with regard to "QUA" or "REN" group membership.

Evaluation of the classification functions associated with the multiple discriminant model showed that almost 80% of the 195 valid cases in the analysis were classified correctly (in terms of their "QUA" or "REN" group membership). The accuracy of classification was particularly poor amongst the true members of the "REN" group: 98% of the "QUA" group and only 26% of the smaller "REN" group were found to have been classified correctly. In other words, the *post hoc* application of the classification functions resulted in the mis-classification of 76% of those members of the BCom cohort who had been excluded from UCT on academic grounds; application of the classification functions to new cases would be likely to result in an even poorer predictive accuracy.

Although it would therefore appear that the multiple discriminant model developed here has little potential in the predictive classification in terms of the potential "QUA" or "REN" group membership of prospective BCom students, multiple discriminant analysis showed that the FACPOINT variable was the most important predictor of the ultimate undergraduate academic outcome of BCom students amongst the list of variables in the analysis. None of the other variables was found to discriminate effectively between the graduates and the academic exclusions within this cohort.

5.2c BCom cohort - Summary

The descriptive analysis of the 1992 BCom cohort revealed that there were substantial differences in the graduation rates within the DET, CO and OMA matriculation authority groups (which were 33%, 56% and 70% respectively). Graduation rates were also seen to increase markedly with increasing faculty points group membership (from 20% amongst the <45 faculty points group to 91% within the 55+ faculty points group). The relatively high graduation rate within the 1992 BCom cohort concealed the fact that only 36% of the cohort had completed their basic degrees within the minimum three year enrolment period. The proportions of students graduating within three years varied markedly as a function of both the matriculation authority group (where none of the DET students, 21% of the CO students and 43% of the OMA students had completed their degrees within three years) and the faculty points group (where the proportion of students graduating within three years increased from 4% of the <45 faculty points group to 69% of the 55+ faculty points group). Chi-square analysis confirmed that there were significant two-way associations between both the matriculation authority group and the faculty points group, and the time to degree within this cohort of BCom students.

The rate of academic exclusion within this cohort (21% overall) was found to be particularly high amongst the DET students (38%) and especially low amongst the OMA students (17%). The rate of academic exclusion was also seen to decline from 56% amongst the <45 faculty points group to only 3% of amongst the 55+ faculty points group. Similarly, the rate of voluntary drop-out (which was 14% for the whole cohort) appeared to be considerably higher within the DET group (29%) than within with either the CO or the OMA group. Voluntary drop-out was also more common amongst the lower faculty points groups: 24% of the <45 faculty points group and only 3% of the 55+ faculty points groups were found to have dropped out of UCT in good academic standing. The largest proportions of the voluntary drop-outs (43%) and of the academic exclusions within this cohort (50%) were found to have occurred following one year of enrolment at UCT.

The best-fit log-linear model describing the three-way matric authority*faculty points group*final undergraduate outcome revealed that there was a significant interaction between the design variable FACGROUP (faculty points group) and the response variable LASTCODE (representing "QUA" or "REN" group membership), and that the graduation:academic exclusion ratio for the cohort as a whole was forced downwards by the relatively high rates of academic exclusion and low rates of graduation within the lower faculty points groups.

The faculty points variable (in the ungrouped format) was also selected as the only significant predictor of the final undergraduate academic outcome (in terms of "QUA" or "REN" group membership within this cohort by means of multiple discriminant analysis. However, the high Wilks' *lambda* observed for the multiple discriminant model combined with the relatively poor predictive powers of the associated classification functions (in terms of "REN" group membership) suggested this model had little potential worth in the prediction of the likely undergraduate academic outcome of prospective BCom students.

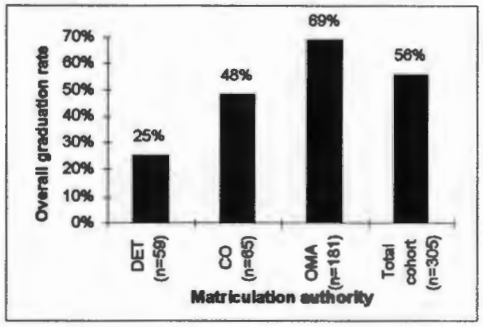
5.3 ANALYSIS OF THE BSc COHORT, FACULTY OF SCIENCE

5.3a Descriptive analysis

A total of 343 FU students entered the BSc degree program within the Science Faculty in 1992. Of these, 38 (11%) subsequently transferred to other faculties at UCT; the majority of these students (87%) transferred out of the faculty following one year of study. The graduation rate amongst the remaining 305 students (henceforward referred to as the BSc cohort) is shown in Exhibit 5.15, which shows that only 56% of this group had graduated by the end of 1996. Graduation rates were seen to vary considerably in relation to matric authority grouping: 25% of the DET students, 48% of the CO students and 69% of the OMA group had successfully completed their BSc degrees. Likewise, there were marked differences in the graduation rates amongst the different faculty point groupings; the lowest

rate of graduation (35%) observed within the 40-44 faculty points group, was in stark contrast to the 84% graduation rate amongst the 55+ faculty points group.

Exhibit 5.15 : Comparison of graduation rates in relation to matriculation authority and faculty points group (BSc cohort)



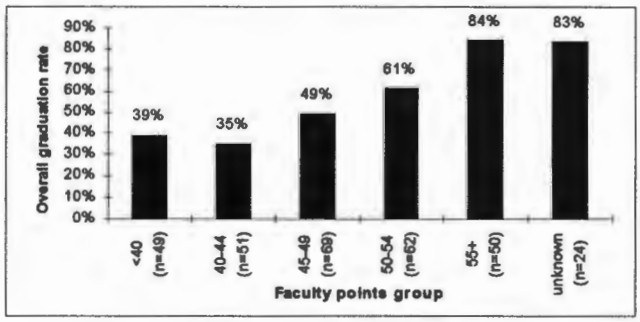


Exhibit 5.16 explores the time to degree amongst graduates of the 1992 cohort in relation to matric authority and faculty points group. Only 36% of the cohort completed their degrees within the minimum 3 year period, and the 3 year degree completion rate was seen to vary markedly as a function of both matric authority and faculty points group: the 3 year degree completion rate increased from 10% amongst the <40 faculty points group to 72% amongst the 55+ group. Whilst 52% of the OMA students were found to have completed their BSc's within 3 years, a considerably smaller proportion of the CO students (29%) and only 2% of the DET students had graduated within 3 years.

Exhibit 5.16 : Time to degree in relation to matriculation authority and faculty points group (BSc cohort)

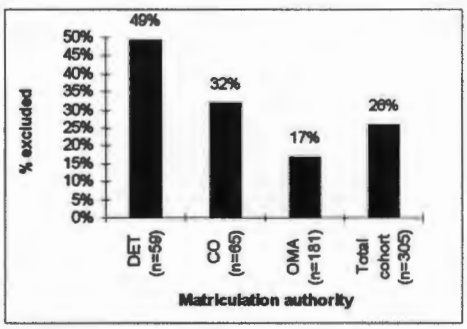
	MATRIC AUTHORITY :			TOTAL COHORT	FACULTY POINTS GROUP :					
	DET	CO	OMA		<40	40-44	45-49	50-54	55+	unknown
No. entering Faculty	**59**	**65**	**181**	**305**	**49**	**51**	**69**	**62**	**50**	**24**
No. graduating after:										
3 years	1	19	94	**114**	5	9	19	31	36	14
	2%	*29%*	*52%*	*37%*	*10%*	*18%*	*28%*	*50%*	*72%*	*58%*
4 years	8	9	27	**44**	9	7	15	5	5	3
	14%	*14%*	*15%*	*14%*	*18%*	*14%*	*22%*	*8%*	*10%*	*13%*
5 years	6	3	4	**13**	5	2	0	2	1	3
	10%	*5%*	*2%*	*4%*	*10%*	*4%*	*0%*	*3%*	*2%*	*13%*
Overall graduation rate	15	31	125	**171**	19	18	34	38	42	20
	25%	*48%*	*69%*	*56%*	*39%*	*35%*	*49%*	*61%*	*84%*	*83%*

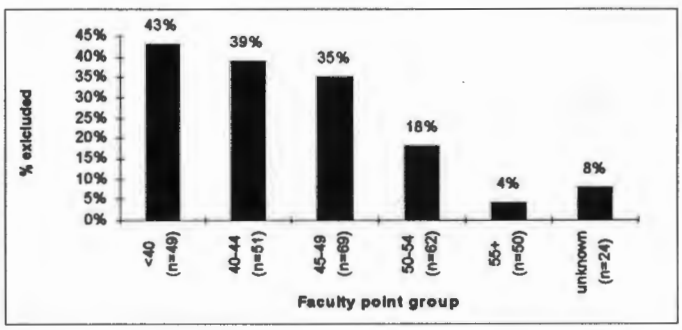
Chi-square analysis indicated that there were highly significant associations between both the matric authority group (Pearson chi-square = 39.073 at df=4, with p=.00000), and the faculty points group (Pearson chi-square = 38.483 at df=8, with p=.00001), and the time to degree (categorised into two groups, *viz.* three years or more than three years) amongst the graduates within this cohort . Time to degree amongst the graduates therefore depended on

matric authority group with comparatively smaller proportions of the DET and CO graduates as opposed to the OMA graduates completing their degrees within three years. Similarly, the faculty points group contributed significantly to the time to degree: smaller proportions of the graduates from within the lower faculty points groups were found to have completed their degrees within the three-year minimum period.

Rates of academic exclusion within the matric authority and faculty points groups, and within the cohort as a whole are illustrated in Exhibit 5.17. This exhibit demonstrates how the low graduation rate within the cohort (56%) was clearly related to the high rate of academic exclusion (26%). Similarly, graduation rates within the different matric authority and faculty points groupings clearly varied inversely with the rates of academic exclusion within each group: the high rate of academic exclusion amongst the DET students (49%) contributed to the low rate of graduation (25%) within this group. The rate of exclusion amongst the CO group (32%) was also markedly higher than that within the OMA group (17%), whilst the graduation rates shown for these two groups (Exhibit 5.15) were 48% and 69% respectively. The rates of academic exclusion amongst the <40 and 40-44 matric points groups (43% and 39% respectively) were seen to exceed the rates of graduation in these groups (39% and 35% respectively); there was a marked drop in the academic exclusion rate amongst the 50-54 faculty point group (18%) in comparison with that amongst the 45-59 point group (35%). Only 4% of the 55+ faculty point group had been excluded from UCT on academic grounds.

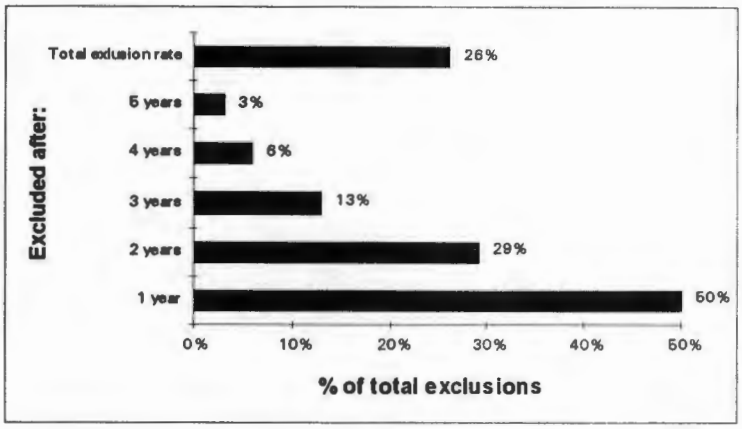
Exhibit 5.17 : Academic exclusion rates in relation to matriculation authority and faculty points group (BSc cohort)





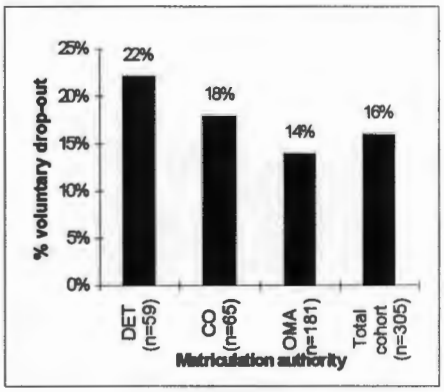
As was the case with the BCom cohort, the uneven distribution of academic exclusions within the Science cohort is alarming as is the fact that half of these exclusions occurred following 2 or more years of enrolment within the Science Faculty (Exhibit 5.18).

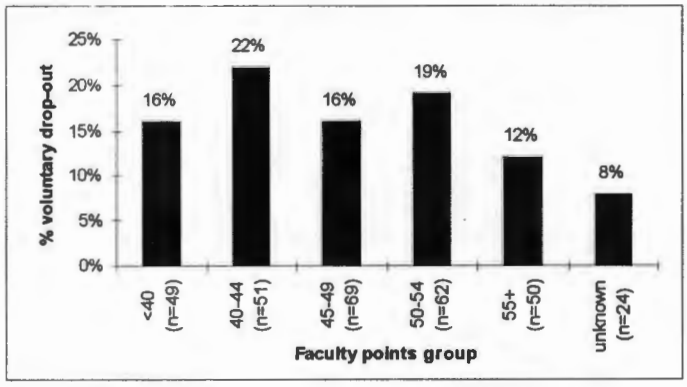
Exhibit 5.18 : Academic exclusions in relation to duration of enrolment (BSc cohort)



Student attrition resulting from voluntary drop-out resulted in the loss of 16% of the 1992 Science cohort (Exhibit 5.19). The rate of voluntary drop-out was highest amongst DET students (22%) and lowest within the OMA matric authority group (14%). However, Exhibit 5.18 indicates that there was no clear trend in the rate of voluntary drop-out as a function of faculty points group; particularly high drop-out rates were observed within the 40-44 and 50-54 faculty points groups (22% and 19% respectively). The lowest drop-out rate (12%) occurred within the 55+ faculty points group where there was also the lowest academic exclusion rate (3%), and thus the highest graduation rate (79%).

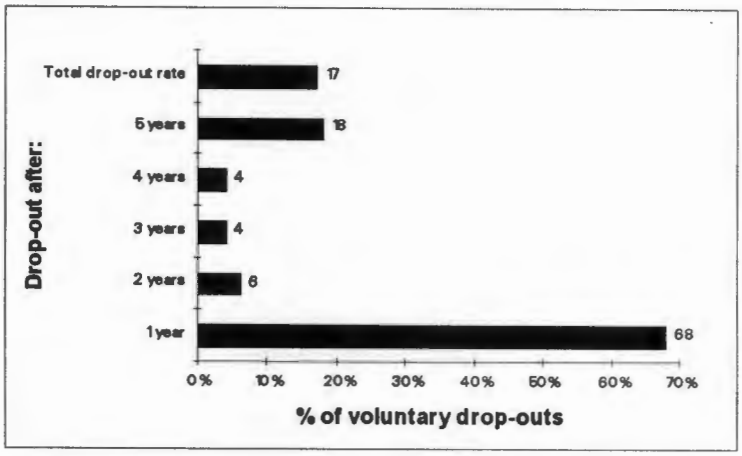
Exhibit 5.19 : Voluntary drop-out rates in relation to matriculation authority and faculty points group (BSc cohort)





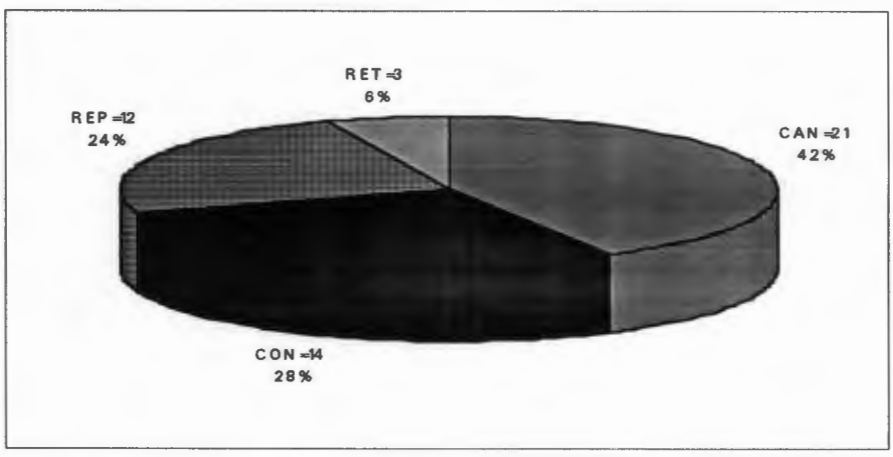
The largest proportion of the voluntary drop-outs (68%) took place at the end of the first year of study (see Exhibit 5.20), but a substantial proportion of these drop-outs (18%) occurred following 5 years of study.

Exhibit 5.20 : Voluntary drop-outs in relation to duration of enrolment (BSc cohort)



The last undergraduate progress codes achieved by the voluntary drop-outs are summarised in Exhibit 5.21. This analysis indicates that voluntary drop-out was possibly related to academic difficulties in only 30% of such cases (i.e. those failing to re-register having achieved a "REP", or a "RET" code). The reasons for failing to re-register amongst 28% of the voluntary drop-outs (whose last undergraduate record reflected the code "CON"), and amongst the 42% of these students who UCT after cancelling their registration (last undergraduate progress code = "CAN"), remain unknown and beg further, more qualitative investigation.

Exhibit 5.21 : Last undergraduate progress codes amongst voluntary drop-outs (BSc cohort)



Only 4 members (1.3%) of the 1992 BSc cohort were still enrolled in the Science Faculty as undergraduates in 1997.

5.3b Statistical analysis

i) Log-linear modelling

Log-linear modelling was employed in order to detect statistically significant interactions between the variables AUTHORN, FACGROUP and LASTCODE within a three-way cross-tabulation representing the 1992 BSc cohort. The development of the log-linear model is described in detail in section 6c of Appendix 1. The numeric coding of the variables in the analysis is explained in Chapter 4.

Chi-square analysis showed that there were highly significant associations between the variables AUTHORN and LASTCODE (chi-square=36.395 with df=2 and p=.0000), and between the variables FACGROUP and LASTCODE (chi-square=33.282 with df=4 and p=.0000). The log-linear analysis returned a best-fit model which incorporated (i) the interaction between the design variables FACGROUP and AUTHORN, (ii) the interaction between the design variable AUTHORN and the response variable LASTCODE, and (iii) the interaction between the design variable FACGROUP and the response variable LASTCODE. The interaction between the design variables resulted from the fact that the faculty points groups were not evenly represented amongst the matriculation authority points groups within the cohort.

Further analysis of the marginal table pertaining to the significant interaction between the design variable FACGROUP and the response variable LASTCODE revealed that there was a progressive increase in the QUA:REN ratio with increasing faculty points groups. The calculated ratios ranged between 0.49 for FACGROUP=1 (i.e. faculty points <40) and 12.43 for FACGROUP=5 (i.e. 55 faculty points or more). The marginal table revealed that the QUA:REN ratio for the whole cohort was 1.85, indicating that the graduation rate for the cohort was forced downwards by the low rates of graduation within the lower faculty points groups.

Similarly, analysis of the marginal table depicting the significant interaction between the design variable AUTHORN and the response variable LASTCODE indicated that the QUA:REN ratios (which are the resultant of the graduation rates academic exclusion rates) varied from between 0.49 for the DET group and 3.66 for the OMA group. The ratio for the CO group (1.34) was intermediate between the ratio for the DET group and that for the OMA group. The graduation:academic exclusion ratio was less than 1 within the DET group only, i.e. the rate of academic exclusion exceeded the graduation rate within this group. The overall graduation:academic exclusion ratio for the cohort as a whole (1.85) was clearly forced downwards by the relatively low ratios for the DET and CO groups.

The results of the log-linear analysis indicated that there was no significant 3-way (i.e. FACGROUP*AUTHORN*LASTCODE) interaction within the 1992 Science cohort. This would

imply that there is no interaction between the FACGROUP*LASTCODE effect and the variable AUTHORN. Similarly, there is no significant interaction between the AUTHORN*LASTCODE effect and the variable FACGROUP.

ii) Multiple discriminant analysis

The development of the multiple discriminant model for the 1992 BSc cohort is described in Appendix 2. Multiple discriminant analysis was carried out in order to select predictors of "QUA" or "REN" group membership within the variable LASTCODE from a list of independent variables. Multiple discriminant analysis was also used in order to derive classification functions which could be used to classify new cases (i.e. new students) as either "QUA" or "REN" group members. Examination of the inter-correlations between the potential predictor variables - SEXN (gender), AUTDUM1 (where the CO matric authority was coded as 1), AUTDUM2 (where the DET matric authority was coded as 1), FACPOINT (the ungrouped faculty points for each student), ENGPTS (the Swedish points equivalent of the matric English symbol) and MATPTS (the Swedish points equivalent of the matric mathematics symbol) - revealed that there was a moderate positive correlation between the variables FACPOINT and MATPTS (R=.60 - see Appendix 2). Experimental exclusion of the variable MATPTS had no effect on the outcome of the model, and thus all of the independent variables were included in the analysis. The numeric coding of the variables in the analysis is explained in Chapter 4.

The multiple discriminant analysis summary indicated that there was a significant discrimination (F=16.295 with p<.000) between the "QUA" and "REN" groups within the ultimate undergraduate academic outcome variable LASTCODE, and that the variables FACPOINT, AUTDUM2 (where the DET matric authority was coded as 1), AUTDUM1 (where the CO matric authority was coded as 1) and SEXN (the gender variable) were all significant predictors of "QUA" or "REN" group membership. However, the moderately high value observed for Wilks' *lambda* for the model (0.77064) suggested that the model possessed only moderate discriminatory power with regard to "QUA" or "REN" group membership. The variable FACPOINT, with the lowest partial *lambda* value, was shown to be the greatest contributor to the prediction of group membership.

Evaluation of the classification functions associated with the multiple discriminant model showed that almost 76% of the 172 valid cases in the analysis were classified correctly (in terms of their "QUA" or "REN" group membership) by means of these functions. The classification was however found to be more accurate amongst the "QUA" cases (89% correct *post hoc* classification) than amongst the "REN" cases (51% correct classification). In other words, the *post hoc* application of the classification functions resulted in the mis-classification of 49% of the members of the BSc cohort who had been excluded from UCT on academic grounds. Although application of the classification functions to new cases would be

likely to produce a poorer predictive accuracy, the classification functions derived here could be usefully applied in determining the likely academic outcomes of potential BSc students, based on the predictor variables detected in this analysis. It would however be advisable to check the predictive capabilities of the classification functions by means of applying them to new cases (with known undergraduate academic outcomes) before applying the functions to potential FU BSc students.

5.3c BSc cohort - Summary

The descriptive analysis of the 1992 BSc cohort revealed that only 56% of the cohort had successfully completed their BSc degrees by the end of the 1996. This analysis showed that there were marked differences in the graduation rates amongst the different matriculation authorities represented within the cohort: particularly small proportions of the DET matriculants (25%) and of the CO matriculants (48%), in comparison with 69% of the OMA students, had graduated within the Faculty by the end of 1996. The proportion of graduates tended to increase amongst increasing faculty points groups (from 39% amongst the <40 points group to 84% amongst the 40+ faculty points group), although there was a particularly low graduation rate (35%) amongst the 40-44 faculty points group.

A relatively small proportion of the 1992 cohort (37%) were found to have graduated within three years of commencing their BSc degrees. There were considerable differences in the three year graduation rates within the different matric authority groups: this analysis showed that 2% of the DET group, 29% of the CO group and 52% of the OMA group had graduated within the minimum three year enrolment period. The analysis also revealed that the three year graduation rate increased with increasing faculty points groups: only 10% of the <40 faculty points group, and 72% of the 55+ faculty points group had completed their degrees within the minimum three year enrolment period. Chi-square analysis revealed that there were highly significant two-way associations between both the matric authority group, and the faculty points group, and the time to degree within this cohort.

The rate of academic exclusion within this cohort (26% overall) was found to be particularly high amongst the DET students (49%); the rate of academic exclusion amongst the CO students (32%) was almost twice as high as that amongst the OMA students (17%). The rate of academic exclusion was seen to decline in relation to increasing faculty points group; this decline was gradual between the <40, 40-44 and 45-49 faculty points groups, but dropped sharply from 35% amongst the 45-49 faculty points group to 18% within the 50-54 faculty points group. Only 4% of 55+ faculty points group were ultimately excluded on academic grounds. Voluntary drop-out was most common amongst the DET students (22% of this group left the University in good academic standing), but there was no clear trend in the rates of voluntary drop-out in relation to faculty points group. The greatest proportions of the both

the voluntary drop-outs (50%) and the academic exclusions within this cohort (61%) were found to have occurred following one year of enrolment at UCT.

Log-linear modelling of the three-way relationship between the matriculation authority, faculty points group and "QUA" or "REN" group membership within the final undergraduate outcome variable revealed that there were significant interactions between both of the design variables and the response variable. Examination of the marginal tables suggested that the graduation:academic exclusion ratio within the cohort was forced downwards by the relatively low graduation rates (and high rates of academic exclusion) within the DET and CO matriculation authority groups, and within the lower faculty points groups.

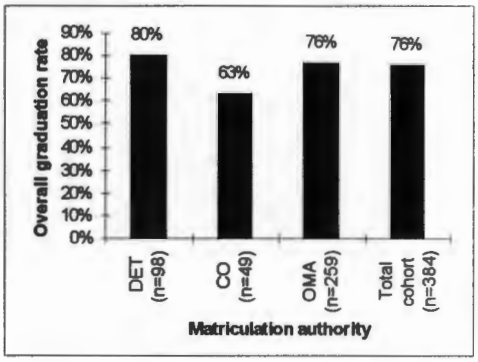
The multiple discriminant analysis carried out in order to detect significant predictors of the final undergraduate academic outcome amongst a range of independent variables confirmed that the ungrouped faculty points variable, matriculation within either the DET or the CO examination authority and the gender variable were all significant predictors of "QUA" or "REN" group membership. The faculty points variable (followed by the DET matriculation authority and then the CO matriculation authority) was found to be the strongest predictor of the final undergraduate academic outcome. The moderately high Wilks' *lambda* for the multiple discriminant model, coupled with the moderate predictive powers of the associated classification functions in terms of "REN" group membership in particular, suggest that the multiple discriminant analysis would be of limited use in predicting the likely group membership of new data cases (e.g. prospective students, or those within other year of entry cohorts).

5.4 ANALYSIS OF THE BA/BSocSc COHORT, FACULTY OF SOCIAL SCIENCE AND HUMANITIES

5.4a Descriptive analysis

In 1992, 426 FU students entered the combined BA and BSocSc programmes in the Faculty of Social Science. Of these, 42 students (10% of the group) transferred to other faculties at UCT: the largest proportion of these transfers (14 or 35%) took place following the first year of study. The graduation rate amongst the remaining 384 students (henceforward referred to as the BA/BSocSc cohort) is illustrated in relation to matric authority and faculty points groups in Exhibit 5.22 below. It appears that the overall graduation rate within the BA/BSocSc group was 76%, and that the highest rate of graduation (80%) occurred amongst the DET students; the overall graduation rate amongst the CO students (63%) was, however, substantially lower than the rate for the whole cohort. Exhibit 5.22 indicates that there was no clear trend in the relationship between the faculty points group and the graduation rate within this cohort, although the lowest overall graduation rate (67%) was observed amongst the highest faculty points group (i.e. those with 40+ faculty points).

Exhibit 5.22 : Comparison of graduation rates in relation to matriculation authority and faculty points group (BA/BSocSc cohort)



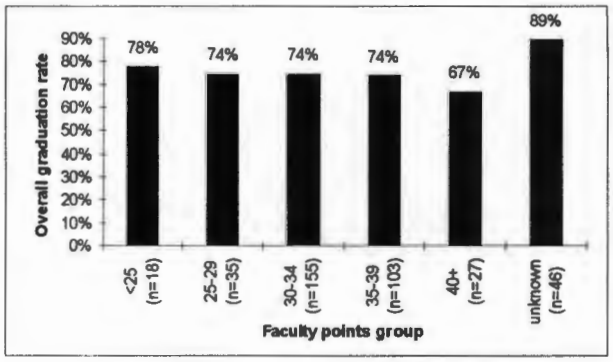


Exhibit 5.23, which summarises the time to degree amongst the various matric authority and faculty points groups, indicates that 54% of the cohort had graduated within the minimum 3 year period. The 3 year graduation rate amongst the OMA students (63%) was considerably higher than that amongst either the DET group (36%) or the CO group (41%). Although the overall graduation rate was highest amongst the DET group, DET students had tended to take longer than their OMA colleagues to complete their basic degrees. Examination of the 3-year graduation rate in relation to increasing faculty points group revealed a general upwards trend from 39% amongst the <25 points group up to a level of 65% amongst the 35-39 points group; the 3-year graduation rate amongst the 40+ faculty points group (52%) was however lower than the overall rate for the cohort as a whole. The proportions of students completing their 3-year degrees within 4 or 5 years therefore declined from 39% of the <25 points group to only 9% of the 35-39 points group, indicating that efficiency (in terms of time to degree) improved with increasing faculty points up to the 35-39 point level.

Exhibit 5.23 : Time to degree as a function of matriculation authority and faculty points group (BA/BSocSc cohort)

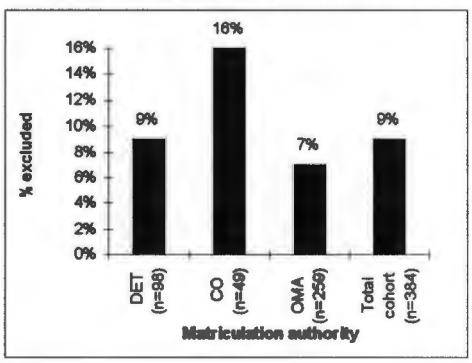
	MATRIC AUTHORITY :			TOTAL COHORT	FACULTY POINTS GROUP :					
	DET	CO	OMA		<25	25-29	30-34	35-39	40+	unknown
No. entering Faculty	**76**	**49**	**259**	**384**	**18**	**35**	**155**	**103**	**27**	**46**
No. graduating after:										
3 years	27	20	162	**209**	7	11	76	67	14	34
	36%	*41%*	*63%*	*54%*	*39%*	*31%*	*49%*	*65%*	*52%*	*74%*
4 years	23	8	31	**62**	2	12	33	8	3	4
	30%	*16%*	*12%*	*16%*	*11%*	*34%*	*21%*	*8%*	*11%*	*9%*
5 years	11	3	5	**19**	5	3	6	1	1	3
	14%	*6%*	*2%*	*5%*	*28%*	*9%*	*4%*	*1%*	*4%*	*7%*
Overall graduation rate	61	31	198	**290**	14	26	115	76	18	41
	80%	*63%*	*76%*	*76%*	*78%*	*74%*	*74%*	*74%*	*67%*	*89%*

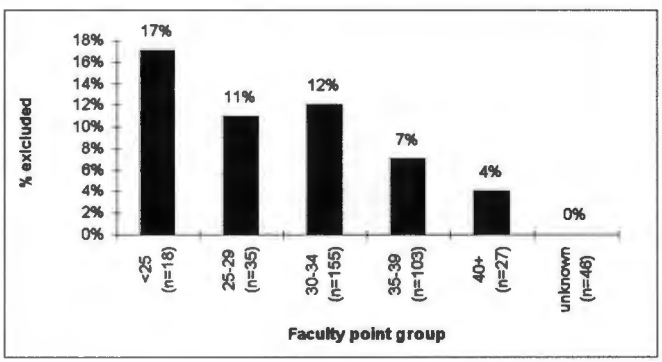
Chi-square analysis revealed that there was a highly significant association between the time to degree (categorised into two groups, *viz.* three years or more than three years) and the

matric authority group amongst the BA/BSocSc cohort graduates (Pearson chi-square = 33.659 at df=2, with p=.00000). In addition, there is a highly significant association between the faculty points group and the time to degree (categorised as above into three year and more than three year groups) amongst the graduates within this cohort (Pearson chi-square = 26.606 at df=4, with p=.00002). Larger proportions of the graduates within the higher faculty points groupings than those in the lower faculty points groups had completed their degrees within three years; similarly, a substantially larger proportion of the graduates within the OMA matric authority group in comparison with the other matric authority groups had completed their degrees within 3 years.

Student attrition rates resulting from academic exclusion are explored in Exhibit 5.24. There was a low overall rate of academic exclusion amongst The BA/BSocSc cohort (9%), but Exhibit 5.24 shows that a particularly large proportion of the CO students (16%) had been lost due to academic exclusion. The rate of academic exclusion generally declined with increasing faculty points grouping (from 17% amongst the <25 points group to only 4% amongst the 40+ group).

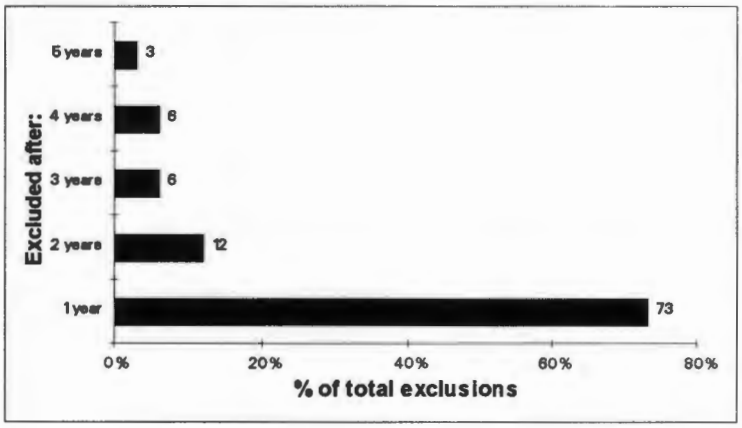
Exhibit 5.24 : Academic exclusion rates as a function of matriculation authority and faculty points group (BA/BSocSc cohort)





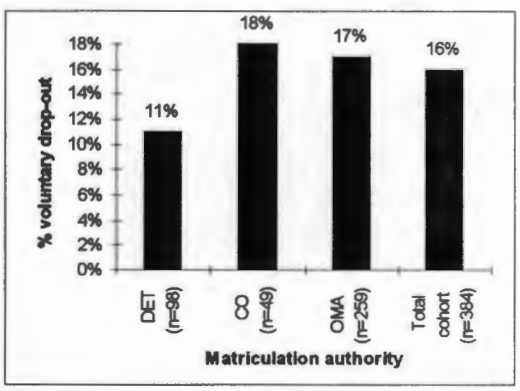
Coupled with the low overall exclusion rate is the observation that the great majority of the academic exclusions (73%) were found to have taken place at the end of the first year of study (see Exhibit 5.25). Inefficiency arising from academic exclusion was therefore minimal within this cohort.

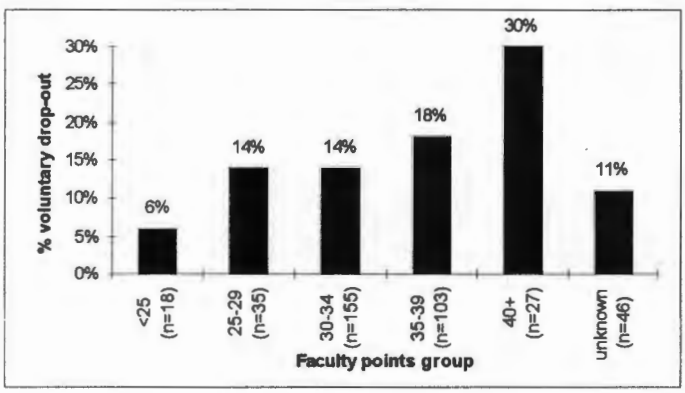
Exhibit 5.25 : Academic exclusions in relation to period of enrolment (BA/BSocSc cohort)



The loss of students due to voluntary drop-out (16% of the BA/BSocSc cohort) was almost twice as large as that due to academic exclusion (9%, see above). Exhibit 5.26 shows that similar proportions of the CO and OMA students (18% and 17% respectively) had left UCT voluntarily, and that the drop-out rate was considerably lower (11%) amongst the DET matric authority group. Interestingly, the rate of voluntary drop-out was also seen to increase with increasing faculty points group from 6% of the <25 points group to 30% of the 40+ points group. The high rate of voluntary drop-out amongst the 40+ points group therefore accounted for the unusually low graduation rate (61%) within this group. The high rate of voluntary drop-out within the BA/BSocSc cohort (relative to the rate of academic exclusion) is cause for concern, and the loss of 30% of the faculty's potentially promising students (in terms of prior academic achievement) is particularly unsatisfactory in terms of potential through-put into postgraduate programs at UCT.

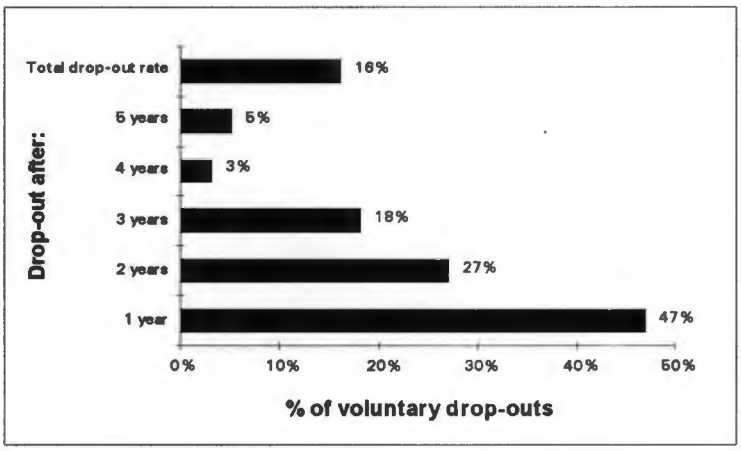
Exhibit 5.26 : Voluntary drop-out rates in relation to matriculation authority and faculty points group (BA/BSocSc cohort)





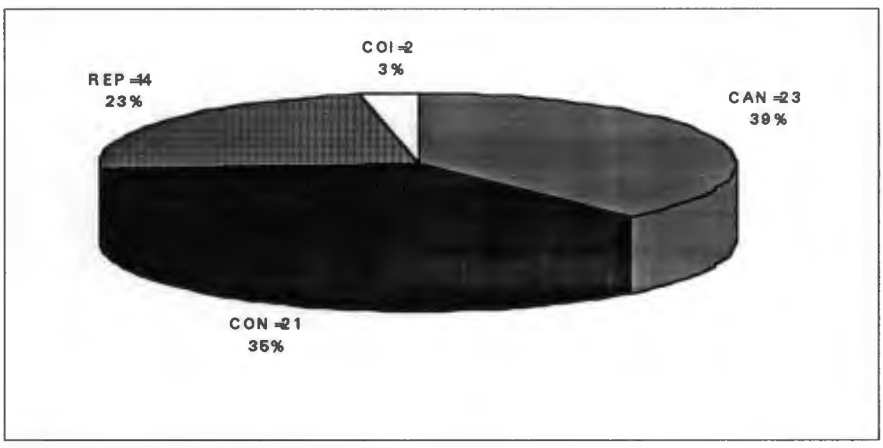
Moreover, the inefficiency due to voluntary drop-out is exacerbated by the observation that 53% of these students dropped out after 2 or more years of study (see Exhibit 5.27).

Exhibit 5.27 : Voluntary drop-outs in relation to duration of enrolment (BA/BSocSc cohort)



A summary of the last academic progress codes amongst the group of voluntary drop-outs is shown in Exhibit 5.28). This summary indicates that academic difficulty may have been a factor in only 26% of these cases (i.e. those whose last undergraduate record reflected a "REP" or a "COI" code), but that other factors (e.g. financial problems or difficulties in social integration) could have led the remainder of the drop-outs to withdraw from UCT (those with progress code "CAN": 39% of the drop-outs) or to fail to re-register at UCT in order to complete their degrees (i.e. those with a final progress code of "CON": 35% of the cases of voluntary drop-out). Again, this situation requires further, more qualitative investigation.

Exhibit 5.28 : Last undergraduate progress codes amongst voluntary drop-outs (BA/BSocSc cohort)



Only 1 member of the 1992 BA/BSocSc cohort (0.26% of the group) was still enrolled for undergraduate studies within the Social Science Faculty in 1997.

5.4b Statistical analysis

i) Log-linear modelling

Log-linear modelling was employed to detect interactions between the design variables FACGROUP and AUTHORN and the response variable LASTCODE. The development of the log-linear model is described in detail in Appendix 1, and the numeric coding of the variables in the analysis is described in Chapter 4. Two-way chi-square analyses revealed that there was no significant association between either the design variable AUTHORN and the response variable LASTCODE (chi-square=5.343 at df=2, with p=.06916), or between the faculty points grouping variable FACGROUP and the response variable LASTCODE (chi-square=2.755 at df=4, with p=.59958). Log-linear analysis returned a best-fit model which incorporated only (i) the interaction between the design variables FACGROUP and AUTHORN and (ii) variable LASTCODE as a main effect. The interaction between the design variables resulted from the fact that the matriculation authority points groups were not evenly represented amongst the faculty points groups within the cohort; students within the lower faculty points groups were largely DET students. The fact that there was no significant interaction between either of the design variables and the response variable LASTCODE implies that the frequencies of graduation and exclusion within the variable LASTCODE were not statistically associated with either of the design variables, and that factors other than those in the log-linear analysis were involved.

ii) Multiple discriminant analysis

Multiple discriminant analysis was carried out firstly to detect predictors of "QUA" or "REN" group membership within the variable LASTCODE from a list of independent variables. A second aim of the multiple discriminant analysis was to derive classification functions which could be used to classify new cases (such as potential new students) as either "QUA" or "REN" group members; such classification functions could be useful in informing admissions policies and in targeting students at risk of academic exclusion for remedial intervention. Prior examination of the inter-correlations between the potential predictor variables - SEXN (gender), AUTDUM1 (where the CO matric authority was coded as 1), AUTDUM2 (where the DET matric authority was coded as 1), FACPOINT (the ungrouped faculty points for each student), ENGPTS (the Swedish points equivalent of the matric English symbol) and MATPTS (the Swedish points equivalent of the matric mathematics symbol) - indicated that there was a moderate positive correlation between the variables FACPOINT and MATPTS (R=.54), as well as a moderate negative correlation between the variables MATPTS and AUTDUM2 (R=-.52). Experimental exclusion of the variable MATPTS however had no effect on the outcome of the multiple discriminant analysis, and all of the predictor variables listed above were therefore included in the analysis the numeric coding of the variables in the analysis is described in Chapter 4. The development of the multiple discriminant model is detailed in Appendix 2.

The multiple discriminant analysis summary indicated that there was a significant discrimination (F=11.097 with p<.0010) between "QUA" and "REN" groups within the ultimate undergraduate academic outcome variable LASTCODE, and that the variable ENGPTS (the Swedish points equivalent of the English matric symbol) was the only independent variable amongst the list of potential predictor variables which discriminated between "QUA" and "REN" group membership within the BA/BSocSc cohort. The high Wilks' *lambda* for the model (0.94958) suggested that the model discriminated weakly between "QUA" or "REN" group membership.

Evaluation of the classification functions associated with the multiple discriminant model confirmed this. Examination of the classification matrix showed that although *post hoc* application of the classification functions resulted in the correct classification (in terms of their "QUA" or "REN" group membership) of 88% of the cases in the analysis, 100% of the "QUA" group and none of the considerably smaller "REN" group were found to have been classified correctly. In other words, applying these classification functions resulted in the mis-classification of 100% of the members of the BA/BSocSc cohort who had been excluded from UCT on academic grounds. Bearing in mind that the classification discussed here is a *post hoc* classification (i.e. a re-classification of the cases from which the classification functions were derived) and that application of the classification functions to new cases would result in an even poorer predictive accuracy, it can be concluded that the multiple discriminant model had little potential in the predictive classification in terms of the potential "QUA" or "REN" group membership of prospective BA/BSocSc students.

5.4c BA/BSocSc cohort - Summary

The descriptive analysis of the 1992 BA/BSocSc cohort revealed that 76% of the cohort had successfully completed a basic degree within the Faculty of Social Science by the end of the 1996 academic year. There were small differences between the overall graduation rates amongst DET students and OMA students (80% and 76% respectively), but the graduation rate amongst the CO matriculants within this cohort (63%) was considerably lower than that observed amongst the other matriculation authority groups. There was no clear trend apparent in the investigation of overall graduation rates in relation to faculty points groups, although a particularly low rate of graduation (67%, in comparison with 76% in the cohort as a whole) was detected within the highest faculty points group.

An examination of the time to degree amongst the graduates within the cohort however revealed that considerably smaller proportions of the DET and CO matriculants (36% and 41% respectively), in comparison with 63% of the OMA matriculants, had completed their degrees within the minimum three-year enrolment period. The three year graduation rate was seen to vary markedly with faculty points group, although particularly small proportions of the 25-29 faculty points group and the 40+ faculty points group (31% an 52% respectively) had

completed their degrees within three years of commencing their studies. Chi-square analysis confirmed that there were highly significant two-way associations between both the matric authority and the faculty points group, and the time to degree within the BA/BSocSc cohort.

The rate of academic exclusion within this cohort (9% overall) was found to be particularly high amongst the CO students (16%) and slightly higher amongst the DET students (9%) than amongst OMA students (7%). The rate of academic exclusion declined (from 17% amongst the <25 faculty points group to only 4% of amongst the 40+ faculty points group. Conversely, the rate of voluntary drop-out within this cohort (16% overall, which was considerably higher than the overall rate of academic exclusion) was seen to be lowest amongst the DET students (11%), and tended to increase with increasing faculty points group from 6% amongst the <25 faculty points group to a level of 30% amongst the 40+ faculty points group. The greatest proportion of the voluntary drop-outs (47%) and of the academic exclusions (58%) were found to have occurred following one year of enrolment at UCT.

Log-linear modelling of the three-way relationship between the matriculation authority, faculty points group and final undergraduate outcome variable revealed that there were no significant interactions between either the matriculation authority group or the faculty points group and the final undergraduate outcome; the only significant interaction within the three-way cross-tabulation was that between the design variables, which resulted from the unequal faculty points group distributions amongst the matriculation authority groups.

The results of a forward, stepwise multiple discriminant analysis carried out in order to detect significant predictors of the final undergraduate academic outcome amongst a range of independent variables confirmed that neither the matriculation authority nor the ungrouped faculty points variable were significant predictors of the final undergraduate academic outcome. Indeed, the Swedish points equivalent of the matriculation English symbol was found to be the only variable tested which discriminated between the groups of graduates and academic exclusions within this cohort. However, the high Wilks' *lambda* for the multiple discriminant model, coupled with the poor predictive accuracy of the classification functions with regard to membership of the group of academic exclusions, indicated that the multiple discriminant analysis would be of little use in predicting the likely group membership of new data cases (e.g. prospective students, or those within other year of entry cohorts).

5.5 INTERFACULTY COMPARISON

A summary of the graduation rates, rates of academic exclusion and voluntary drop-out within the four cohorts examined is presented in Exhibit 5.29 below:

Exhibit 5.29: Comparison of the graduation rates, rates of academic exclusion and rates of voluntary drop-out within the four selected cohorts

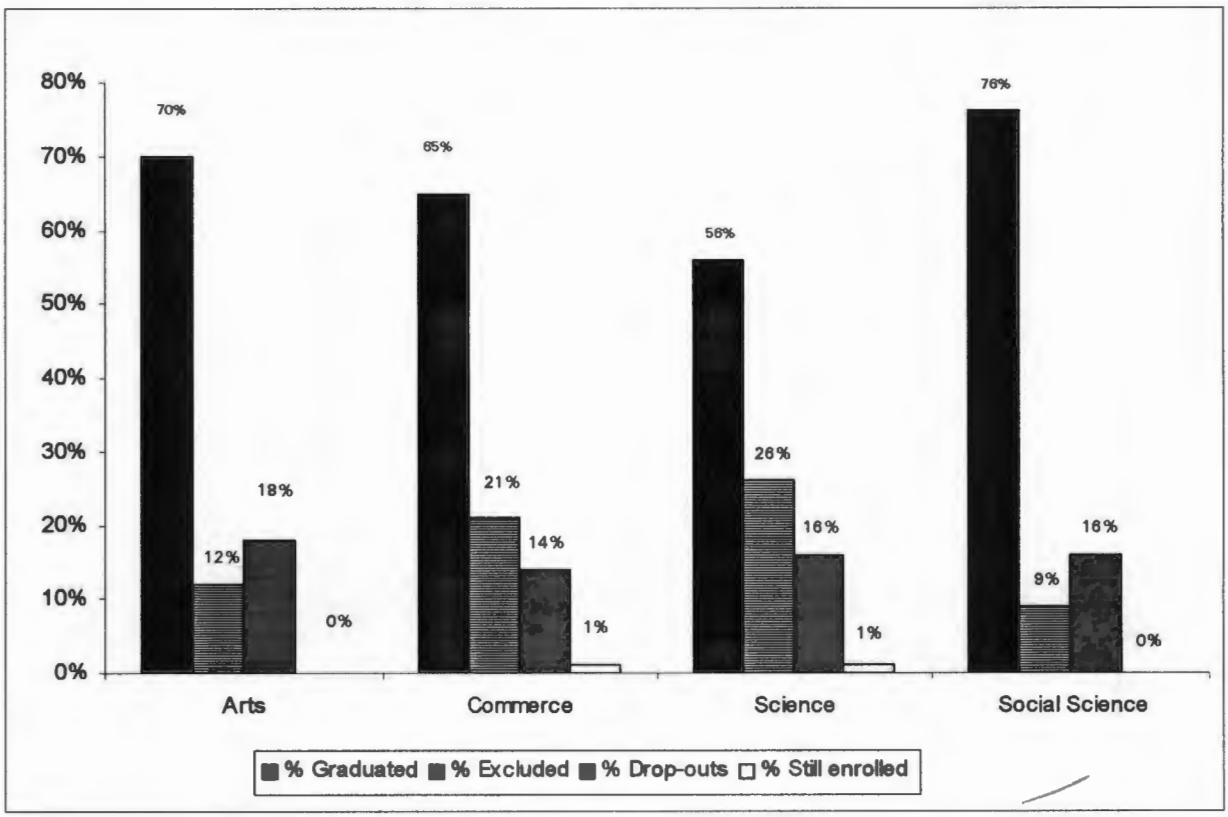


Exhibit 5.29 indicates that the voluntary exclusion rate did not vary greatly amongst the four cohorts, but that the marked differences in the overall graduation rates (which ranged from 56% within the Science cohort to 76% within the Social Science cohort) were complemented by widely varying rates of academic exclusion (from 9% within the Social Science cohort to 26% within the Science cohort). The more nuanced analysis of the individual cohorts presented above explores the added dimensions of time and heterogeneity in terms of prior educational achievement within the cohorts, and therefore highlights particular areas of concern within the different cohorts. The problem areas noted above include:

- the relatively low proportions of students completing their 3-year degrees within the 3 year enrolment period; this problem is particularly pronounced amongst the DET and CO matric authority groups, and within the lower faculty point groups in each cohort;
- the comparatively low graduation rates amongst the DET students in the BCom and BSc cohorts; in both cohorts the rate of exclusion amongst the DET students was in excess of the graduation rate, and voluntary drop-out rates within this group were also particularly high;

- the relatively low graduation rates amongst CO students in the Arts and Social Science cohorts; within these cohorts, the graduation rates within the CO matric authority groups were lower than those within the DET group. In all four cohorts the graduation rate amongst CO students was considerably lower than that amongst OMA students;

- an examination of the last undergraduate progress codes amongst the voluntary drop-out component within each cohort exposed the academic diversity of the group; large proportions of the voluntary drop-outs and had failed to re-register at UCT despite having achieved a "CON" progress code, or had cancelled their registration at the University and not returned. The quantitative nature of the cohort analysis presented above does little to shed light on the reasons for voluntary drop-out from UCT, but is does serve to highlight the alarmingly high rates of voluntary drop-out within the 35-39 faculty points group in the Arts cohort and within the 40+ faculty points group within the Social Science cohort.

The results of the descriptive cohort analyses are discussed in relation to their internal efficiency implications in Chapter 6. A discussion of the outcomes of the log-linear models and of the multiple discriminant analyses described above, and their implications of these findings with regard to admissions and enrolment management during the transformation process at UCT, is also presented in Chapter 6. Possible reasons for the poor discriminatory powers and predictive powers of the multiple discriminant models are also debated.

CHAPTER 6

DISCUSSION OF RESULTS

The research described in this report aimed primarily to quantify student progression through the four large bachelors' degree programmes at UCT in relation to prior matriculation performance. In this way, student throughput and graduation within these programmes were investigated in order to detect where high levels of repetition and attrition were compromising the internal efficiencies in the production of graduates. The secondary aim of this research was to explore the development of predictive models of student retention within the selected cohorts which could be used in the formulation of admissions policies and enrolment management strategies. The research results pertaining to these two key aims are discussed below in terms of:

- graduation rates and time to degree;
- students attrition (due to academic exclusion and voluntary drop-out);
- the duration of enrolment prior to academic exclusion or voluntary drop-out; and
- the development of predictive models of student retention within the selected degree programmes.

6.1 GRADUATION RATES AND TIME TO DEGREE

The cohort survival analysis of the four cohorts of first-time entering bachelor's students revealed that there were marked differences in the graduation rates within the different cohorts: by the end of 1996, 76% of the BA/BSocSc (Faculty of Social Science and Humanities) cohort, 70% of the BA (Faculty of Arts) cohort, 65% of the BCom (Faculty of Commerce) cohort and only 56% of the BSc (Faculty of Science) cohort had successfully completed their basic bachelors' degrees. The production of graduates was therefore considerably more efficient within the BA and BA/BSocSc programmes than within the BCom, and particularly within the BSc programme. The cohort survival analyses revealed that there were small differences in the rates of voluntary drop-out amongst the four cohorts: the voluntary drop out rate was lowest within the BCom cohort (14%), was intermediate (16%) within the both the BSc and BA/BSocSc cohorts, and was highest (19%) within the BA cohort. Substantial differences in the rates of academic exclusion within the four cohorts were however observed: there were particularly low rates of academic exclusion within the BA/BSocSc and BA cohorts (9% and 12% respectively), and considerably higher rates of academic exclusion within the BCom and BSc cohorts (21% and 26% respectively). The observed variation in efficiencies with regard to the "production of graduates" between the four cohorts had therefore largely resulted from differences in the rates of academic exclusion within the cohorts.

In the interests of optimising the internal efficiency in the production of graduates, an important consideration is the proportion of each cohort successfully completing their basic degrees within the minimum three year enrolment period. A comparison of the four cohorts examined in this work reveals a clear dichotomy between the BA and BA/BSocSc cohorts (within which 53% and 54% of each cohort respectively graduated within the minimum three year period) and the BCom and BSc cohorts where 36% and 37% respectively graduated within three years of their initial enrolment. There was thus considerable scope for the improvement of internal efficiency by means of increasing the proportions of students completing their degrees within the minimum three year period within all of the cohorts examined, but this was particularly true of BCom and BSc cohorts.

6.2 STUDENT ATTRITION

6.2a Student attrition due to academic exclusion

Although there were no clear trends in the overall graduation rates as in relation to increasing faculty points groups within the BA and BSocSc cohorts (where the absence of a clear trend in the rate of voluntary drop-out resulted in a similar absence of a pattern in graduation rates in relation to increasing faculty points groups), there were marked patterns in the rates of academic exclusion in relation to faculty points group within all four cohorts: in each case there was a clear decline in the rate of academic exclusion depicted as a function of increasing faculty points group. The extremely high rates of academic exclusion amongst the <45 faculty points group within the BCom cohort (56%), and amongst the <40 faculty points group within the BSc cohort (43%) are particularly alarming. The debate around the formulation of admissions policies for BSc and BCom students should take cognisance of these figures which suggest that the present levels of academic intervention are not providing effective remediation for students entering these programmes with particularly low faculty points ratings.

Whilst UCT recognises that "...DET students in particular have generally been seriously under-prepared for traditional university study" (UCT, 1994:15), the cohort survival analysis performed here has indicated that within two of the cohorts, students entering UCT through the Coloured Senior Certificate (referred to here as members of the CO group of matriculants) did not perform as well as those from the DET (Department of Education and Training) schools. Within the BA cohort, for example, the lowest graduation rate (58%) and highest rate of academic exclusion (31%) occurred amongst this group; the proportion of the CO matriculants graduating within the minimum three year enrolment period (38%) was also the lowest within the cohort. Similarly, within the BA/BSocSc cohort, the rate of academic exclusion amongst the CO students (16%) was the highest within the cohort, and the graduation rate within this group (63%) was considerably lower than that amongst the DET

students (80%). However, 41% of the CO students, in comparison with 36% of the DET students, were found to have graduated within three years of entering UCT, indicating that the rate of repetition was higher amongst the ex-DET students within this cohort.

By contrast, students within the DET matriculation authority group appeared to be the least prepared for tertiary level studies within the BCom and BSc cohorts: only 33% of the DET students within the BCom cohort and 25% of those within the BSc cohort had successfully completed their basic degrees by the end of 1996. The rates of academic exclusion amongst these two groups were particularly high (38% within the BCom cohort and 49% within the BSc cohort), as were the voluntary drop-out rates (29% amongst the DET students within the BCom cohort and 22% amongst the DET students within the BSc cohort). Within both the BCom and the BSc cohorts, therefore, the great majority of the ex-DET students failed to complete their basic degrees at UCT, and the "revolving door" metaphor mentioned in the White Paper on higher education transformation (Department of Education, 1997:12) could be applied to these students.

It would therefore appear to be of great importance in terms of internal institutional efficiency that subsequent cohorts should be tracked through the University in order to substantiate or to repudiate the attrition patterns seen within the 1992 FU cohorts. Moreover, it would seem that a need exists within the Commerce and Science faculties in particular to re-evaluate their current efforts in providing academic remediation to students from disadvantaged educational backgrounds.

6.2b Student attrition due to voluntary drop-out

Although the rates of voluntary drop-out did not vary greatly amongst the four cohorts analysed (18% of the BA cohort, 16% of both the BSc and BA/BSocSc cohorts, and 12% of the BCom cohort were found to have left UCT in good academic standing), the drop-out patterns observed within the four cohorts were quite different. Within the BA cohort, voluntary drop-out was most common amongst students from the previously white matriculation authorities (20% amongst the OMA students); examination of the voluntary drop-out rate as a function of faculty points group revealed that this phenomenon was most common amongst the second highest faculty points group (i.e. the 35-39 points group). The analysis of the BA/BSocSc cohort showed that the rates of voluntary drop-out were similar amongst the CO matriculation authority group (18%) and amongst those from the previously white matriculation authorities (17%), and that voluntary drop-out was most common amongst the 35-39 and 40+ faculty points groups (18% and 30% of these groups respectively). Examination of the last undergraduate progress codes of the voluntary drop-outs revealed that the great majority of the voluntary drop-outs (56% of those within the BA cohort and 74% of those with the BA/BSocSc cohort) had either cancelled their registration at UCT, or had

failed to re-register for undergraduate studies at UCT having achieved a favourable academic progress code. The cohort survival analysis therefore indicated that factors other than academic difficulty appeared to have prompted the greater proportion of the voluntary withdrawals.

An examination of the rate of voluntary drop-out in relation to matriculation authority and faculty points group within the BCom cohort clearly showed that the drop-out rate was highest amongst the ex-DET students (29% of these students), and declined with increasing faculty points group. Examination of the final undergraduate progress codes amongst this group nevertheless showed that the majority of these students (57%) had either cancelled their registration or had failed to re-register following the successful completion of a particular academic year. Within the BSc cohort, the largest voluntary drop-out rate was observed amongst the DET matriculants, but there was no apparent pattern in the relationship between voluntary drop-out rates and increasing faculty points groups. Moreover, 70% of the drop-outs within the BSc cohort had either cancelled their registration or had failed to re-register having achieved a "CON" academic progress code (*i.e.* following a successful year of study). The analysis of the BSc and BCom cohorts therefore also indicated that factors other than academic difficulty had prompted the majority of the voluntary withdrawals.

The cohort survival analysis presented here has thus served to quantify voluntary drop-out within the four selected cohorts, and has established that the voluntary drop-out phenomenon within the four cohorts examined encompassed a diversity of academic outcomes. In view of the relationships demonstrated in this analysis between faculty points groups and the rate of voluntary drop-out rate, it would appear that both the Arts and the Social Science faculties in particular were losing predominantly "good" students (as measured in terms of their matriculation performance) due to voluntary drop-out. Because no attempts have been made thus far to follow up voluntary drop-outs, either telephonically or by means of postal surveys, in order to establish why they decided to leave UCT, the reasons for these losses remain unclear. The analysis of the selected cohorts has therefore served to stress the need for a more qualitative investigation of what motivates students to leave UCT, and to identify the barriers to degree completion perceived by these students. It is possible that the findings of this type of qualitative analysis with regard to students' reasons for leaving UCT (possibly based on an evaluation of both the external and internal factors discussed in Chapter 2) could serve as a basis for improving institutional student retention, thereby improving efficiency in the production of graduates.

6.3 THE DURATION OF ENROLMENT PRIOR TO ACADEMIC EXCLUSION OR VOLUNTARY DROP-OUT

The simplified cohort survival analysis described in this report has therefore clearly identified problems with regard to the rates of both academic exclusion and voluntary drop-out within the four cohorts analysed. The results of the analyses have also quantified the rates of both forms of student attrition in relation to the duration of enrolment. If we regard the expenditure on students who leave the University without completing their basic degrees as "wasted expenditure", then the timing of both forms of student attrition in relation to the duration of enrolment has clear efficiency implications: student attrition becomes increasingly costly to the University (and, of course, to the students involved) as the period of enrolment prior to attrition increases. Expressed in simple terms, students who are excluded on academic grounds or who withdraw voluntarily from UCT following two or more years of enrolment constitute a greater "wastage" of University resources than those who leave or who are excluded after one year of study.

The results of the analysis show that within all four cohorts, at least half of the academic exclusions (58% of those within the both the BA cohort and BA/BSocSc cohorts, and 50% of those within the both the BCom and the BSc cohorts) occurred at the end of 1992, i.e. following only one year of study. Conversely, 50% of the academic exclusions within both the BCom and BSc cohorts took place after two or more years of enrolment at UCT; smaller, but still substantial proportions of the academic exclusions within the BA and BA/BSocSc cohorts took place following two or more years of enrolment. A review of the findings of the cohort analyses with regard to the timing of the voluntary withdrawals from the selected cohorts shows that more than half of the withdrawals from both the BCom and the BA/BSocSc cohorts took place following two or more years of study at UCT. Although the majority of the voluntary withdrawals from both the BA and the BSc cohorts (61% of these withdrawals in both cohorts) took place at the end of the 1992 academic year, substantial proportions of those students who left UCT in good academic standing were found to have done so after two or more years of enrolment at the University. This synopsis of the relationship between student attrition (due to both academic exclusion and voluntary withdrawal) and the duration of enrolment prior to attrition suggests that there is a significant level of "wasteful" expenditure on students who leave the University after two or more years of enrolment without completing a basic degree. Consequently, it would appear that an explicit institutional focus on retaining students with an enrolment duration of more than one year could produce substantial gains in internal efficiency.

6.4 THE DEVELOPMENT OF PREDICTIVE MODELS OF STUDENT RETENTION WITHIN THE SELECTED DEGREE PROGRAMMES

6.4a Log-linear modelling

The descriptive analyses presented in this report exposed a number of problem areas and differentials with regard to student progression measured in relation to the matriculation authority and faculty points groupings within each of the cohorts. Having explored the patterns in the graduation and academic exclusion rates within the selected cohorts, and having established that the phenomenon of voluntary drop-out within each cohort encompassed a diversity of academic outcomes (which would therefore not be predictable in terms of prior educational achievement), log-linear modelling was employed in order to explore the relationships between the graduate ("QUA") and academic exclusion ("REN") outcomes within each cohort, in relation to the key matriculation criteria of matriculation authority and faculty points group. Log-linear modelling was therefore used as a means of exploring the interactions within three-way cross-tabulations of the matriculation authority group, the faculty points group and ultimate undergraduate outcome in terms of membership of the group of graduates or membership of the group of academic exclusions.

Using log-linear modelling, a significant interaction between the grouped matriculation authority and faculty points variables (which constitute the design variables in these analyses) was detected within each of the selected cohorts. This interaction indicated that CO matriculants and particularly DET matriculants within the cohorts tended to be concentrated within the lower faculty points groups, a situation may have arisen due to UCT's stated commitment to redressing the inequities of the South African schooling system by means of admitting applicants "from those sectors of the education system which ...have not provided adequate opportunities for students to demonstrate their ability but who are assessed as having the ability to succeed in a particular academic programme" (UCT, 1994:9). In all four cohorts, the interaction between the grouped matriculation authority and faculty points variables was found to be an essential component of the log-linear model describing the data frequencies within the three-way cross-tabulations.

Although the importance of the design variable interaction was common to each of the cohorts analysed, log-linear modelling detected notable differences in the interactions between the design variables and the two-level ("QUA" or "REN") response variable amongst the selected cohorts. Within the BA cohort, log-linear modelling detected a significant interaction between the matriculation authority group and the final academic outcome, indicating that there were statistically different frequencies of graduation and academic exclusion within the three matriculation authority groups. The analysis however showed that there was no significant association between the faculty points group and the final

undergraduate academic outcome as defined by "QUA" or "REN" group membership. Conversely, within the BCom cohort, log-linear modelling detected a significant interaction between the faculty points group and the frequencies of graduation and academic exclusion, but showed that there was no significant interaction between the matriculation authority group and the final undergraduate outcome within this cohort. Within the BSc cohort, both the matriculation authority group and the faculty points group were found to be associated (separately, but not jointly) with the final undergraduate outcome measured in terms of "QUA" or "REN" group membership. Neither the matriculation authority group nor the faculty points group was found to be associated with the final undergraduate academic outcome within the BA/SocSc cohort; this implies that factors other than the design variables in the analysis had influenced the frequency distribution of the response variable data within the three-way matriculation authority*faculty points group*final undergraduate outcome table.

It is important to note that log-linear modelling did not detect any significant three-way interactions (*i.e.* involving both of the design variables and the response variable) within any of the four selected cohorts. This would imply that where the faculty points group was found to be associated with the final undergraduate outcome (*i.e.* within the BCom and BSc cohorts), this interaction was not affected by the students' matriculation authority group membership; in other words the graduation rate within these cohorts improved with increasing faculty points groups regardless of the students' matriculation authority group membership. Conversely, where matriculation authority grouping was found to interact significantly with the final undergraduate academic outcome (within the BA and BSc cohorts), this effect was not modified by the matriculation authority group membership of the students; CO students within these cohorts had therefore performed particularly poorly regardless of their prior educational achievement (measure in terms of faculty points).

In summary, log-linear analysis revealed that the final undergraduate outcome as measured in terms of "QUA" or "REN" group membership was associated with the matriculation authority of members of the BA cohort, with the faculty points group of members of the BCom cohort, and an additive (but not interactive) combination of the matriculation authority and the faculty points group of members of the BSc cohort. Log-linear analysis showed that there were no significant interactions between the final undergraduate outcome and either the matriculation authority or the faculty points groups within the BA/BSocSc cohort (within the Faculty of Social Science), suggesting that factors other than those tested in this analysis had influenced student progression within this cohort.

6.4b Multiple discriminant analysis

Having explored the relationships between the categorical matriculation authority, the faculty points group and the final undergraduate outcome variables within the four selected cohorts,

multiple discriminant analysis was employed in order to detect significant discriminators of "QUA" or "REN" group membership amongst a range of potential (ratio-scaled) predictor variables. The multiple discriminant analyses also aimed to formulate classification functions by means of which the likely graduates and academic exclusions amongst new cases (such as prospective students) within each cohort could be determined, based on the significant predictor variables detected in the analysis. The formulation of satisfactory classification functions would enable admissions planners to predict (amongst a group of applicants to a particular programme) the likely graduates and the probable academic exclusions, assuming the absence of academic interventions over and above those already in place during the 1992 -1996 period. A second assumption underlying this classification would be that, in view of the academic diversity amongst the voluntary drop-outs described above, individuals identified as either potential graduates or potential academic exclusions by means of the classification functions could ultimately become voluntary drop-outs for a range of possible (and as yet unknown) reasons.

The multiple discriminant analysis of the BA cohort revealed that both the ungrouped faculty points variable and membership of the CO matriculation authority were significant predictors of the final undergraduate academic outcome within this cohort: CO matriculants with low faculty points were classified as likely academic exclusions in terms of the multiple discriminant model. Within the BCom cohort, the ungrouped faculty points variable alone was found to discriminate significantly between the graduates and the academic exclusions; testing of the classification functions revealed that students with less than 44 faculty points would be classified as academic exclusions under the multiple discriminant model. Within the BA/BSocSc cohort, the numerical (Swedish points) equivalent of the matric English symbol was found to be the only significant predictor of graduation or academic exclusion; students with Swedish points equivalents for matriculation of 4 (which is the Swedish points equivalent of a higher grade "E" symbol, or a standard grade "C" symbol) or more were classified as graduates by means of the classification functions derived under the multiple discriminant model. Examination of the overall values for Wilks' *lambda* for each of these cohorts however indicated that these multiple discriminant models possessed only weak discriminatory powers with regard to the separation of graduates and students excluded on academic grounds. This conclusion was supported by the evaluation of the classification functions derived in terms of the multiple discriminant models: within all three cohorts, the *post hoc* classification of students who were ultimately excluded on academic grounds was found to be highly inaccurate, although the graduates amongst each group were most often classified correctly.

The application of multiple discriminant analysis to the BSc cohort revealed that a number of the independent variables in the analysis discriminated effectively between the graduates and the academic exclusions within this cohort. The ungrouped faculty points variable, DET

matriculation authority, CO matriculation authority and gender variables were all found to discriminate effectively between the graduates and the academic exclusions within the BSc cohort. On the basis of the partial Wilks' *lambda*, the ungrouped faculty points variable was found to be the strongest predictor of graduation or academic exclusion. Inspection of the overall Wilks' *lambda* indicated that the model possessed moderate discriminatory powers with regard to the graduates and the academic exclusions amongst the cohort. A *post hoc* classification of the cases within the BSc cohort confirmed this perception: although the majority of the successful graduates within the cohort were classified correctly by means of the classification functions, the proportion of correct classifications amongst those who had been excluded on academic grounds was just over 50%.

By means of multiple discriminant analysis, key predictor variables of the final undergraduate academic outcome within each of the four cohorts were therefore identified. In most instances, the significant predictor variables identified by means of multiple discriminant analysis substantiated the interactions demonstrated by means of log-linear modelling. Within the BA cohort however, multiple discriminant analysis identified the ungrouped faculty points variable as a second significant predictor of graduation or academic exclusion, whereas log-linear analysis failed to detect a significant association between the grouped faculty points variable and the ultimate undergraduate academic outcome. This would suggest that there was some loss of predictive accuracy in the grouping of the faculty points variable into pentile (five point) ranges. It is possible that a different grouping system, such as a more sensitive re-grouping of the 30-34 faculty points group which contained the largest proportion of the cases, may have resulted in the identification of a significant association between the grouped faculty points variable and "QUA" or "REN" group membership within the last undergraduate progress code by means of log-linear analysis.

The high values for Wilks' *lambda* observed for the multiple discriminant analyses pertaining to the BA, BCom and BA/BSocSc cohorts indicated that these models discriminated weakly between the graduates and academic exclusions within the selected cohorts. Similarly, the *post hoc* application of the classification functions associated with the multiple discriminant models revealed a poor predictive accuracy with regard to the classification of cases (*i.e.* students) as probable graduates or probable academic exclusions. This poor predictive accuracy was largely associated with the classification of those students within each cohort who had been excluded on academic grounds, a prediction which is fundamental in the formulation of admissions policies and in the identification of those students likely to be excluded on academic grounds in the absence of additional remedial intervention. In view of the likelihood that the application of the classification functions derived in these analyses to new cases (such as prospective students) would result in an even poorer predictive accuracy, it would seem that the multiple discriminant models developed for the BA, BCom and

BA/BSocSc cohorts were of limited value in the context of predicting final undergraduate academic outcomes of prospective students.

Wilks' *lambda* for the multiple discriminant model derived for the BSc cohort was found to be moderately high, and the *post hoc* application of the classification functions associated with this model revealed a moderate predictive accuracy, particularly in terms of the classification of the academic exclusions within the cohort. The application of multiple discriminant analysis to the BSc cohort therefore served to identify significant predictors of "QUA" or "REN" group membership, and also yielded a set of classification functions which could be used (with moderate success) in the classification of prospective students as likely graduates or as likely academic exclusions.

It is possible that the reasons for the relatively weak discriminatory powers and poor predictive capabilities of the multiple discriminant models derived in these analyses may relate to one or more of the following methodological features of the analyses:

- In view of the statistical associations demonstrated between faculty points groupings and the time to degree amongst the graduates in each cohort, the combination of all graduates within each cohort within a single academic outcome group (i.e. the "QUA" group) may be insufficiently sensitive to the nuances within these cohorts.

- Another possibility is that, in view of the particularly poor classification accuracy with regard to the academic exclusions within each cohort (the majority of whom were classified as likely graduates by means of the classification functions), currently unexplored factors within the UCT environment (rather than those measuring prior educational attainment), were instrumental in the final undergraduate academic outcomes of students within these cohorts. Badenhorst *et al* (1990:44), for example, point out that purely quantitative analyses could obscure the effects of a range of non-academic factors including the "subtle racism, feelings of alienation, socio-political influences and concrete problems regarding finances, transport and accommodation" experienced by black students at a largely white university.

- A third possible explanation for the relatively poor predictive capabilities of the classification functions associated with the multiple discriminant models is provided by Mitchell *et al* (1997:383) who point out that, in view of the fundamental differences between school education and university education, a relatively weak correlation between matriculation performance and university achievement should be anticipated within student progression research. However, the relationships demonstrated by these authors between the different matriculation examinations and failure rates at university (Mitchell *et al*, 1997:384) suggest that heterogeneity with regard to the predictive abilities of the

matriculation examinations (particularly within the range of examinations encompassed by the "OMA" group of matriculants) may have had marked effects on the discriminatory powers of the multiple discriminant analyses carried out as part of the present investigation. The differences in the predictive value of the various South African matriculation examinations explored by Mitchell *et al* (1997:387), and their conclusion that the current "overemphasis on matriculation performance is unwarranted", coupled with the poor predictive capabilities of the classification functions derived within the present research all suggest that there is a strong case for the establishment of a national university entrance examination or postsecondary admissions similar to that used within American and Australian postsecondary systems.

The conclusions drawn from the results discussed here, in relation to both the transformation initiatives of the White Paper (Department of National Education:1997) and the institutional financial issues explored in Chapter 3, are presented in Chapter 7 which follows.

CHAPTER 7

CONCLUSIONS

The 1997 White Paper on higher education (Department of National Education, 1997) outlines a set of initiatives which are intended to transform South African tertiary education into a system which is both appropriate to the new South African society, and one which will be responsive to national needs, realities and opportunities. Within the White Paper, there is a strong emphasis on the dual goals of equity of access to higher education and equity in terms of the opportunity to succeed (i.e. equity of outcomes) within the postsecondary system. It is envisaged that these objectives will be achieved by means of controlled expansion of the system, in conjunction with the accelerated provision of bridging programmes and the development of comprehensive student support systems in order to ensure the equitable accommodation of under-prepared black students within the tertiary system. The White Paper stresses that expansion of access to higher education should not lead to a "revolving door" syndrome (Department of National Education, 1997:2) with high rates of failure and drop-out. The commitment to the goal of equity of outcomes is made explicit in the linking of earmarked public funding for equity and redress to "measurable progress toward improving quality and reducing high drop-out and repetition rates" (Department of National Education, 1997:12).

It would therefore appear that South African universities will be required to respond to the dual pressures of expanding access to students from disadvantaged educational backgrounds, and of ensuring equity of outcomes (measured by an increase in graduation rates, and a reduction in repetition and drop-out rates) throughout the student body. In terms of the national policy of fiscal discipline, and in view of the proposed changes to the proportions of formula and non-formula public funds allocations within the tertiary sector (see Chapter 1), it seems likely that historically white institutions such as UCT will have to address these issues against a background of diminished government funding.

As has been pointed out in Chapter 3, there are limits to the extent to which the already apparent decline in government funding can be balanced with income from either increased tuition fees or from the greater diversification of funding sources. A possible further decrease in government funding in real terms could therefore translate into a real decrease in the institutional resources available for the achievement of the transformation objectives set out in the White Paper on higher education transformation. These pressures will compel universities such as UCT to use the financial resources available to them as efficiently as possible. Chapter 1 describes how the expansion of access to under-prepared students will limit the extent to which internal efficiency gains can be made by means of increasing student:staff ratios or by means of cutting back on non-educational expenditure (such as

subsidised student services), and hence how the focus is thus likely to fall on gains in internal efficiency resulting from improved student progression and graduation rates.

In summary, therefore, it would appear that (i) the greater focus on internal institutional efficiency outlined in the White Paper on higher education transformation; and (ii) the stated intention in the White Paper to monitor institutional student progression as a means of ensuring "equity of outcomes", and hence the likelihood that such figures will become part of the required three-year rolling plan; and (iii) the probable financial restraints within which the dual goals of equity of access and equity of outcomes will be required to be accomplished, will result in a greater institutional focus on potential gains in internal efficiency within UCT. It is therefore likely that the University will be compelled to identify possible means of improving institutional internal efficiency, particularly with regard to addressing factors related to student progression, which drive up the costs per graduate.

The overview of some recent gross efficiency indicators presented in Chapter 3 suggested that a decrease in real terms in the total funds available to the University was already in evidence during the 1988 - 1995 period, and that this had been coupled with a small but real increase in the total institutional expenditure, resulting in an increase in UCT's total expenditure relative to the total funds available. Significantly, the real decrease in funds available (which resulted largely from a real decrease in income from government appropriations) coincided with a 12% growth in student headcount enrolments. By the end of 1995, UCT was clearly in the position of having to teach more students using relatively smaller financial resources, whilst simultaneously striving to redress racial and gender inequities in terms of access to the University. The decline seen in the undergraduate throughput rate (from 27% in 1993 to 21% in 1995), and the observation (in terms of a cost unit evaluation) that the production of graduates at the University had become relatively less efficient during the 1993 and 1995 period, may have been related to these increased pressures on University resources.

There are indications, therefore, that even prior to the implementation of the transformation initiatives set out in the White Paper, the costs per graduate at UCT were increasing. In view of the financial and equity considerations set out above, it would seem imperative that the University should address the problem of reducing the inefficiencies underlying the increasing costs per graduate, and it is thus necessary to firstly determine precisely where these inefficiencies lie. Although UCT collects on an annual basis information on indicators such as undergraduate success rates (which are measured as the percentage of courses taken that are passed, and are therefore not indicative of students' overall academic performance) and annual rates of graduation, academic exclusion, voluntary drop-outs (by population group, gender and entrance category), there has until now been no attempt to systematically track homogenous cohorts (in terms of year of entry and entrance category) of

students in order to quantify either the graduation rate and time to degree within such groups, or the rate and timing of academic exclusion and voluntary drop-out.

A basic objective of this research was therefore to gather, by means of cohort survival analysis, this academic progression data in relation to participants in the four large three-year bachelors' degree programmes at UCT. It was envisaged that the data generated for the four selected cohorts could form the foundation for ongoing research in the area of student progression at UCT. This would permit the production of a system of benchmarks against which the performance of successive cohorts could be measured, and in terms of which admissions policies and enrolment plans could be revised on an incremental basis. A secondary objective of this research was to explore the use of two statistical techniques in devising predictive models of final undergraduate outcomes in terms of these matriculation criteria, which could also potentially inform efficiency-related student admission and institutional enrolment planning.

The analyses were confined to first-time entering undergraduate students, thereby excluding those transferring from other faculties at UCT or from other postsecondary institutions as well as those repeating their first year of study, all of whom would differ from first-time entering students in terms of their prior exposure to the tertiary learning environment. The cohort survival analysis focused on each student's final undergraduate academic outcome, and on the role of key matriculation criteria (chiefly the matriculation authority and faculty points scored in the matriculation examinations) in this outcome. In view of the World Bank's (1994:19-20) identification of low graduation rates, high repetition rates and high drop-out rates (including both academic exclusion and voluntary drop-out) as key factors which tend to drive up the costs per graduate, this analysis aimed to provide an initial measure of these important efficiency indicators within each of the four cohorts examined.

The setting of target graduation, repetition and drop-out rates at UCT would clearly be facilitated by the availability of a complete breakdown of the costs of repetition and attrition (including voluntary drop-out and academic exclusions) to both the students involved and to the University. Whilst the direct costs to the student could be measured in terms of tuition and residence fees, the cost of books (and other educational materials) and the loss of income because of their student status, the exact extent of institutional expenditure on students who drop-out in good academic standing or who are excluded on academic grounds without completing their basic degrees has not been quantified. This expenditure could be considered to be "wasted expenditure" which is fundamental to driving up the costs per graduate. In the absence of such financial data, the measurement of student attrition and repetition rates provides a valuable indication of the internal institutional efficiencies in the production of graduates.

The investigation of student progression through the four selected bachelors' degree programmes at UCT by means of cohort survival analysis clearly identified a number of areas within which the institutional efficiencies related to student throughput and graduation rates were unsatisfactory. In particular, low graduation rates (complemented by high attrition rates) were observed amongst students entering the University with low faculty points ratings within all of the cohorts apart from the BA/BSocSc cohort. Similarly, the rates of attrition amongst students from disadvantaged educational backgrounds, i.e. those who had matriculated within the old DET schools and those who had completed the Coloured Senior Certificate (CSC), were found to be markedly higher than amongst those who had matriculated within the formerly white matriculation authorities. The results of these analyses, which confirmed that the graduation rates amongst the CSC matriculants were poorer than those amongst the DET matriculants within both the BA and BA/BSocSc cohorts, are fundamental to the debate around the identification of educational disadvantage amongst current and prospective students at the University of Cape Town. These results also suggest that the academic development programmes in operation within the University during the period of this study (i.e. between 1992 and 1996) were not sufficiently successful in the remediation of students entering UCT from educationally disadvantaged backgrounds.

In addition to identifying problems with regard to the overall graduation rates amongst particular groups of matriculants within each of the cohorts, the cohort survival analyses showed that particularly small proportions of the students (notably within the BCom and BSc cohorts) had completed their studies within the minimum three year period of enrolment. In addition, it was noted that a great deal of the student attrition within each of the cohorts had taken place following an enrolment duration of two or more years. These observations suggest that substantial gains in internal efficiency could be achieved by increasing the proportions of students graduating within three years, and by adopting a strategy of increased retention of students with an enrolment duration of two or more years. A fundamental part of this strategy would be the elucidation (probably by means of a qualitative survey) of the factors leading either to students' ultimate exclusion on academic grounds, or to the personal decision to leave UCT without completing a basic degree.

The statistical analyses performed as part of this study successfully identified significant predictor variables of the final undergraduate outcome within each of the cohorts which, in most instances, confirmed the trends detected in the descriptive analyses of the cohorts. Importantly, the multiple discriminant analysis carried out amongst the BA/BSocSc cohort confirmed the findings of both the descriptive and the log-linear analyses of this cohort, which indicated that neither the matriculation authority nor the faculty points score were associated with the final undergraduate academic outcome within this cohort; this analysis suggested rather that the matriculation English symbol was a key predictor of the final academic outcome within the BA/BSocSc cohort. Although the classification functions

derived in relation to the multiple discriminate models were of limited value with regard to the categorisation of both current and prospective students as likely graduates or as potential academic exclusions, the predictor variables identified by means of these analyses could be of some use in the formulation of admissions policies and in the development of systems designed to detect (for the implementation of remedial intervention) students potentially at risk of academic exclusion. The statistically poor discriminatory powers of the multiple discriminant models developed in relation to the selected cohorts nevertheless indicated that the variables selected for analysis did not adequately describe the variability in the final undergraduate outcomes of student within these cohorts. It is therefore possible that additional quantitative indicators should be included within the predictive models, or that qualitative information derived (for example) from qualitative student exit surveys could yield key discriminatory variables with regard to undergraduate student progression within these cohorts.

In conclusion, it seems that the focus on institutional internal efficiency will intensify in view of the financial constraints within which the transformation objectives outlined in the White Paper on higher education transformation (Department of Education, 1997) are likely to be tackled. In order to improve the financial efficiencies involved in the production of graduates, it is necessary at the outset to identify key areas of inefficiency in terms of student progression. The results of the cohort survival analyses reported here have served to quantify the components of student progression, thereby identifying specific problem areas within each of the selected bachelors' programmes. Although the application of statistical modelling techniques within this study failed to yield useful predictive models of undergraduate academic outcomes in relation to matriculation criteria, these models did detect significant predictor variables of student progression within the selected cohorts. The systematic production of student progression data of the type generated in this analysis, possibly extended to include additional qualitative variables, could provide vital management information for the efficient functioning of the University of Cape Town during the process of transformation.

REFERENCES CONSULTED

1. Anderson, D.R., Sweeney, D.J. & Williams, T.A . (1993) Statistics for Business and Economics. Minneapolis: West Publishing Company.

2. University transformation: Search for cultural synthesis. Financial Mail 02/02/1996, pp.24-26.

3. Badenhorst, F.D., Foster, D.H., & Lea, S.J. (1990) Factors affecting academic performance in first-year psychology at the University of Cape Town. South African Journal of Higher Education, Vol. 4, No. 1, pp.39-45.

4. Bean, J.P. (1990a) Why students leave: insights from research. In Hossler, D., Bean, J.P. and Associates (Eds) The Strategic Management of College Enrollments. San Francisco: Jossey-Bass.

5. Bean, J.P. (1990b) Using retention research in enrollment management. In Hossler, D., Bean, J.P. and Associates (Eds) The Strategic Management of College Enrollments. San Francisco: Jossey-Bass.

6. Bunting, I.A. (1997a) Performance indicators for South Africa's education systems. Social Dynamics, Vol. 23, pp.139-153.

7. Bunting, I.A. (1997b) Universities and Technikons in the Western Cape : 1988 - 1993. Report prepared for National Information and Performance Indicator Project of Educational Policy Unit of University of the Western Cape.

8. Committee of University Principals (1987) Macro-aspects of the university within the context of tertiary education in the Republic of South Africa. Report of the Main Committee of the CUP, Pretoria.

9. Conway, C. (1996) A Preliminary view of undergraduate student attrition at UVic. http://www.inst.uvic.ca/iaviews/volume1/issue3/iaviews3.html.

10. Day, J.H. (1997) Enrollment forecasting and revenue implications for private colleges and universities. In D.T. (Ed) New Directions for Institutional Research. Forecasting and Managing Enrollment and Revenue: An Overview of current Trends, Issues, and Methods. San Francisco: Jossey-Bass.

11. Department of Education (1997) A programme for the transformation of higher education. Education White Paper 3, Pretoria.http://www.polity.org.za/govdocs/white_papers/highed.html

12. EduSource (1997) EduSource Data News, No. 18, October 1997.

13. Everett, J.E. & Robins, J. (1991) Tertiary entrance predictors of first-year university performance. Australian Journal of Education, Vol. 35, No. 1, pp.24-40.

14. Ewell, T.E., Parker, R. & Jones, D.P. (1988) Establishing a Longitudinal Student Tracking System: An Implementation Handbook. Boulder: National Center for Higher Education Management Systems.

15. File, J. (1988) Equal access. Leadership South Africa, Vol. 7, No. 4, pp.107-108.

16. Grobbelaar, J. (1995) The universities as part of the higher education system of South Africa. A.C.U. Bulletin of Current Documentation, No. 119.

17. Hamilton, D.J. (1990) Multiple regression analysis and prediction of GPA upon degree completion. College Student Journal, Vol. 24, pp.91-96.

18. Jawitz, J. & Scott, L. (1997) Who does not succeed in engineering at the University of Cape Town? What can one tell from the retention rate? Centre for Research in Engineering Education (CREE), Research Report 1.

19. Levin, J. & Wyckoff, J. (1988) Effective advising: identifying students most likely to succeed in Engineering. Engineering Education, Vol. 78, pp.178-182.

20. Little, T.M. & Hills, F.J. (1972) Statistical Methods in Agricultural Research. University of California, Davis, California.

21. Mathiasen, R.E. (1984) Predicting college academic achievement: a research review. College Student Journal, Vol. 18, pp.380-386.

22. McDonald, D. (1995) Enhancing the Recruitment and Retention of First-Year Student with Improved Career Selection Opportunities. http://fairway.ecn.purdue.edu/fre/asee/fie95/3a3/3a31/3a31.html

23. Mitchell, G., Fridjhon, P. & Haupt, J. (1997) On the relationship between the matriculation examination and university performance in South Africa - 1980 to 1991. South African Journal of Science, Vol. 93, pp.382-387.

24. Murray, J. (1993) Why students withdraw from university: a case study. Canadian Institutional Researchers and Planners Conference. March 1993. Simon Fraser University, Vancouver.

25. NCHE (National Commission on Higher Education) (1996a) A framework for transformation. Report of the National Commission on Higher Education. Pretoria: CTP Book Printers.

26. NCHE (National Commission on Higher Education) (1996B) Higher Education Financing in South Africa. Report of the Finance Task Group.http://star.hsrc.ac.za/nche/final/tg4/2.html and http://star.hsrc.ac.za/nche/final/tg4/4.html

27. Neely, R. (1977) Discriminant analysis for prediction of college graduation. Educational and Psychological Measurement, Vol. 37 (Winter), pp.965-970.

28. Noblitt, A.T. (1990) Case study: how information systems support enrolment management. In Hossler, D., Bean, J.P. and Associates (Eds) The Strategic Management of College Enrollments.

29. Nora, A. (1990) Campus-based aid programs as determinants of retention among Hispanic community college students, Journal of Higher Education, Vol. 61, pp.312-331.

30. O'Connell, B. (1994) Equity: a case study of the Peninsula Technikon. In Fiske, I. & Sutherland, C. (Eds) Changing by Degrees? Cape Town: UCT Press.

31. Pascarella, E.T., Terenzini, P.T. (1979) Student-faculty contact and college persistence: a further investigation. Journal of Educational Research, Vol. 49 (Spring), pp.214-218.

32. Porter, J.D. & Gebel, M.A. (1995) Arizona State University: Student tracking in a university setting. In Ewell, P.T. (Ed) New Directions for Institutional Research. Student Tracking: New Techniques, New Demands. San Francisco: Jossey-Bass.

33. Ramphele, M. (1994) Equity: reality factors and their implications. In Fiske, I. & Sutherland, C. (Eds) Changing by Degrees? Cape Town: UCT Press.

34. Reinecke, C.J. (1993) The present subsidy formula and the subsidies generated by the humanities and natural sciences. Africa 2001 Symposium and Workshop: The Employability of South African Graduates, Vol. 2, No. 1, pp.16-18. Pretoria: HSRC Printers.

35. Sandel, R.H. & Sydow, D. (1997) Advances - and a retreat - in student retention. Inquiry, Vol. 1, No. 1, pp.61-70.

36. Sarkar, G. (1993) SIASST Retention study: factors affecting retention of first-year students in a Canadian Technical Institute of Applied Science and Technology. Canadian Institutional Researchers and Planners Conference. March 1993. Simon Fraser University, Vancouver.

37. Simkins, C. (1994) Equity, efficiency and tertiary education: the economic context. In Fiske, I. & Sutherland, C. (Eds) Changing by Degrees? Cape Town: UCT Press.

38. Smith, M. (1992) DET Matric: what does it predict? Comment, Vol. 8, Autumn 1992, pp.10-17.

39. St. Louis Community College (1996) Enrollment Management Plan: Student Retention Plan. http://fv.cc.mo.us/admin/emp/G.html.

40. StatSoft Inc (1995) Statistica for Windows [Computer program manual]. Tulsa: StatSoft Inc.

41. Stoker, D.J., Engelbrecht, C.S., Crowther, N.A.S., du Toit, S.H.C. and Herbst, A. (1985) Investigation Into Differential Entrance Requirements For Tertiary Educational Institutions. Pretoria: HSRC.

42. Tinto, V. (1975) Dropout from higher education: a theoretical synthesis of recent research. Review of Educational Research, Vol. 45, Winter, pp.89-125

43. University of Cape Town (1994) The Transformation of the University of Cape Town 1984-1994: A Decade of Change and Development.

44. Van Wyk, J.A. & Crawford, J.L. (1984) Correlation between matric symbols and marks obtained in a first year ancillary physics course at the University of the Witwatersrand. South African Journal of Science, Vol. 80, pp. 8-9.

45. Voorhees, R.A. (1987) Toward building models of community college persistence: A logit analysis. Research in Higher Education, Vol. 26, pp.115-129.

46. Weidman, J.C. (1985) Retention of nontraditional students in postsecondary education. Paper presented at the 69th Annual Meeting of the American Educational Research Association, Chicago. ERIC Document Reproduction Service No.. ED 261 195).

47. Whitaker, D.G. (1987) Persistence and the two-year college student. Paper read at The Association for the Study of Higher Education, Baltimore.

48. Wilson, B.P. (1993) From institutional research to policy implementation: a case study. Canadian Institutional Researchers and Planners Conference. March 1993. Simon Fraser University, Vancouver.

49. Woolston, D.C., Shook, K. and Wilson, J. (1995) Same Problem, Different Solutions: Attempts at Improving Retention in Engineering at a Research vs. a Teaching University. http://fairway/ecn.purdue.edu/fre/asee/fie95/3a3/3a31/3a31.html

50. World Bank (1994) Development in Practice: Higher Education - The Lessons of Experience. Washington: World Bank.

APPPENDICES

APPENDIX 1

LOG-LINEAR ANALYSIS OF RELATIONSHIPS BETWEEN MATRICULATION PERFORMANCE AND UNDERGRADUATE ACADEMIC OUTCOMES.

The development of the log-linear models for each of the selected cohorts is discussed below. The method used in the development of these models is described in Chapter 4 of this report. In all of the analyses, the design variables selected were the matriculation (or matric) authority grouping variable (AUTHORN) and the grouped faculty points variable (FACGROUP); the contents of these variables and their coding are described in Chapter 4 of this report. The response variable in all of the analyses was the final undergraduate outcome variable LASTCODE; only the "QUA" and "REN" codes within the variable LASTCODE were included in the analyses.

FACULTY OF ARTS (BA cohort)

Chi-square analysis was initially employed in order to establish whether or not there were statistically significant associations between either of the design variables (AUTHORN and FACGROUP) and the two-level response variable LASTCODE within the BA cohort. A summary of the results of the chi-square testing is presented in Exhibit 8.1.

Exhibit 8.1 : Summary of Chi-square analyses

Design variable	Pearson Chi-square	df	p	Significant?
AUTHORN	11.117	df=2	P=.00386	YES
FACGROUP	6.927	df=4	p=.13981	NO

The data presented in Exhibit 8.1 indicate a statistically significant association between the matric authority grouping variable (AUTHORN) and the two-level response variable LASTCODE. There was, however, no significant association between the faculty points group variable (FACGROUP) and the response variable (LASTCODE) at the 5% alpha level, indicating that the FACGROUP*LASTCODE interaction would be unlikely to form part of the log-linear model. The results of these analyses showed that there were significant differences in the frequencies of graduation and academic exclusion amongst the three matric authority groupings, but not amongst the faculty point groupings within the Arts cohort.

Examination of all k-factor interactions (Exhibit 8.2) revealed significant two-way interactions, but no significant three-way interactions, between the three variables. Inspection of the tests of marginal and partial association (Exhibit 8.3) indicated that the interactions between variables 1 and 2 (AUTHORN and FACGROUP) and between variables 1 and 3 (AUTHORN

and LASTCODE) were significant, but that there was no significant interaction between variables 2 and 3 (FACGROUP and LASTCODE).

Exhibit 8.2 : Results of Fitting all K-Factor Interactions (art92fu.sta)

	Degrs.of Freedom	Max.Lik. Chi-squ.	Probab. p	Pearson Chi-squ	Probab. p	Significant?
1 (AUTHORN)	7	170.0276	.000000	291.9605	0.000000	YES
2 (FACGROUP)	14	90.1103	.000000	94.5353	.000000	YES
3 (LASTCODE)	8	4.5875	.800604	6.0723	.639138	NO

Exhibit 8.3 : Tests of Marginal and Partial Association (art92fu.sta)

	Degrs.of Freedom	Prt.Ass. Chi-sqr.	Prt.Ass. p	Mrg.Ass. Chi-sqr.	Mrg.Ass. p
1 (AUTHORN)	2	55.23932	.000000	55.23932	.000000
2 (FACGROUP)	4	45.97287	.000000	45.97287	.000000
3 (LASTCODE)	1	68.81525	.000000	68.81525	.000000
12 (AUTHORN*FACGROUP)	8	74.30122	.000000	76.79661	.000000
13 (AUTHORN*LASTCODE)	2	7.04552	.029529	9.54091	.008483
23 (FACGROUP*LASTCODE)	4	3.77280	.437641	6.26808	.180031

On the basis of the evaluation of k-factor interactions and tests of marginal and partial associations, it would seem that the best-fit model would incorporate the 21 (FACGROUP*AUTHORN) and 31 (LASTCODE*AUTHORN) interactions. Automatic fitting of the best-fit model confirmed this; analysis of the goodness of fit of this model (shown in Exhibit 8.4) showed that there was no significant difference between the expected frequencies under the best-fit model and the observed frequencies in the three-way table.

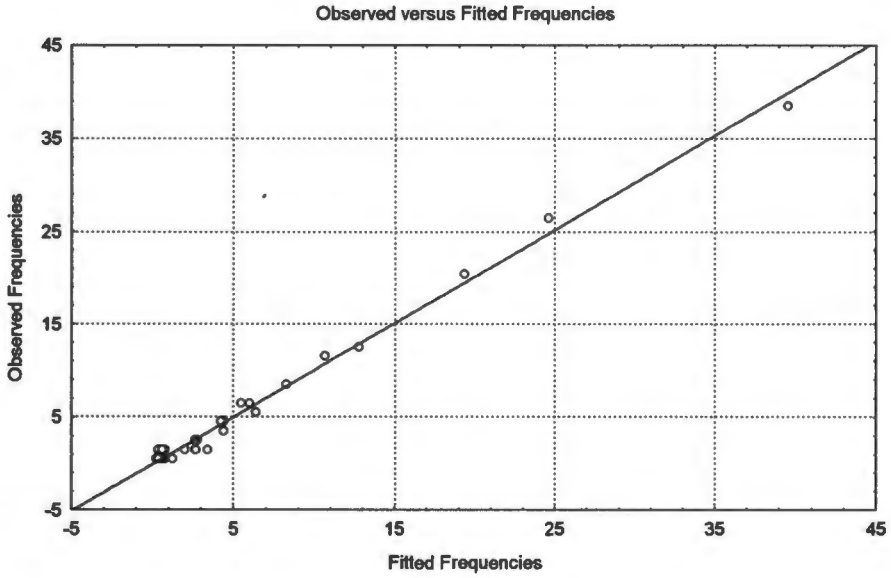
Exhibit 8.4 : Test of goodness of fit of best-fit model

Observed Table: AUTHORN(3) * FACGROUP(5) * LASTCODE(2) (art92fu.sta)
Model: 21 (FACGROUP*AUTHORN), 31 (LASTCODE*AUTHORN)

	Chi-sqr	df	p
Max Likelihood Chi-square	8.360317	12	.756360
Pearson Chi-square	9.963224	12	.619187

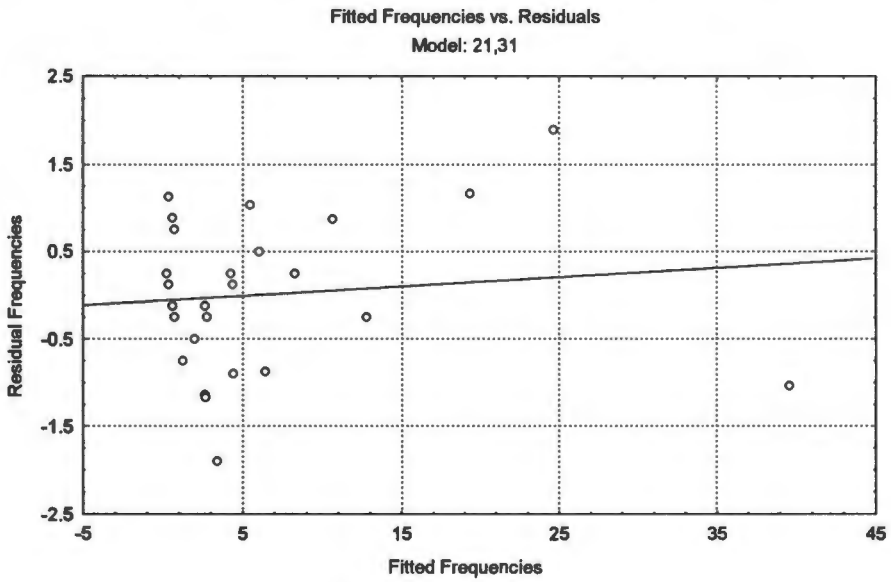
An examination of a plot of the observed frequencies against the fitted frequencies under the best-fit model (Exhibit 8.5) showed a close linear fit, and that there were no major outliers:

Exhibit 8.5 : Plot of observed frequencies vs. fitted frequencies under best-fit model



Furthermore, a plot (Exhibit 8.6) of the fitted frequencies against the residual frequencies (i.e. observed frequencies minus expected frequencies) under the best-fit model produced the random scatter consistent with an appropriate model :

Exhibit 8.6 : Plot of fitted frequencies vs. residual frequencies under the best-fit model



Prior to interpretation of the best-fit model, the effect of omitting the 21 (FACGROUP*AUTHORN) interaction from the model was evaluated by performing an hierarchical test to examine the goodness of fit of model incorporating the 31 (LASTCODE*FACGROUP) interaction only. Exhibit 8.7 below shows that there was a highly

significant difference between the observed cell frequencies and those fitted under the 2,31 model, indicating that the association between the design variables was essential within the best-fit model.

Exhibit 8.7: Test of goodness of fit of model excluding the interaction between design variables

Observed Table: AUTHORN(3) * FACGROUP(5) * LASTCODE(2) (art92fu.sta)
Model: 2,31

	Chi-sqr	df	p
Max Likelihood Chi-square	85.15691	20	.000000
Pearson Chi-square	87.28336	20	.000000

The best-fit model determined by both automatic and manual model fitting did not include the 32 (LASTCODE*FACGROUP) interaction. A second hierarchical test was carried out in order to compare the improvement of fit achieved in including this interaction (i.e. the 21 31 23 model - see Exhibit 8.8 below). The chi-square difference (=3.771, df=4) was not significant (p>.05), confirming that there was no significant interaction between the FACGROUP and LASTCODE variables.

Exhibit 8.8 : Test of goodness of fit of model including FACGROUP*LASTCODE interaction

Observed Table: AUTHORN(3) * FACGROUP(5) * LASTCODE(2) (art92fu.sta)
Model: 21,31,32

	Chi-sqr	df	p
Max Likelihood Chi-square	4.583222	8	.801040
Pearson Chi-square	6.179791	8	.627102

Interpretation of the model was carried out by examining the marginal tables pertaining to the best-fit model. Whilst the marginal table between the design variables AUTHORN and FACGROUP (Exhibit 8.9) indicated that there were relatively more low faculty points DET students than within the other 2 matric authority groups, the nature of the effects depicted in the LASTCODE*AUTHORN (Exhibit 8.10) are of particular interest:

Exhibit 8.9 : Marginal table of AUTHORN*FACGROUP

Marg. Tabl. (freq+delta):AUTHORN by FACGROUP (art92fu.sta)

	AUTHORN DET	AUTHORN CO	AUTHORN OMA	Total
1	11.00000	2.00000	3.0000	16.0000
2	17.00000	1.00000	5.0000	23.0000
3	8.00000	17.00000	45.0000	70.0000
4	1.00000	7.00000	22.0000	30.0000
5	1.00000	1.00000	28.0000	30.0000
Total	38.00000	28.00000	103.0000	169.0000

Exhibit 8.10 : Marginal table of AUTHORN*LASTCODE

Marg. Tabl. (freq+delta):AUTHORN by LASTCODE (art92fu.sta)

	AUTHORN DET	AUTHORN CO	AUTHORN OMA	Total
QUA	28.50000	17.50000	90.5000	136.5000
REN	9.50000	10.50000	12.5000	32.5000
Total	38.00000	28.00000	103.0000	169.0000

Calculation of the QUA:REN ratios for the 3 matric authority groupings from the marginal Table (Exhibit 8.10) indicated that the graduation:exclusion ratio amongst the OMA students (7.24) was considerably higher than that amongst DET students (where the ratio was 3.0), and amongst CO students (QUA:REN ratio = 1.7). The QUA:REN ratio for the whole cohort (4.2) was therefore markedly lower than that amongst OMA students, and was clearly forced down by the lower QUA:REN ratios amongst the DET and particularly the CO students.

FACULTY OF COMMERCE (BCom cohort)

Chi-square analysis was initially employed in order to establish whether or not there were significant two-way associations between either of the design variables (AUTHORN or FACPOINTS) and the two-level response variable LASTCODE. A summary of the results of the chi-square testing (presented in Exhibit 8.11) showed that there were significant associations between both of the design variables and the response variable LASTCODE. In other words, the frequencies of the graduation and of academic exclusion differed significantly amongst both the matric authority and the faculty points groupings.

Exhibit 8.11 : Summary of Chi-square analyses

Design variable	Pearson Chi-square	df	p	Significant?
AUTHORN	10.350	df=2	p=.00566	YES
FACGROUP	33.236	df=3	p=.00000	YES

The results of fitting all K-factor interactions (Exhibit 8.12) indicated that there were significant two-way interactions in addition to the main effects in the model, but that fitting the three-way interaction would not improve the fit of the model. Evaluation of the tests of marginal and partial associations (Exhibit 8.13) showed that the interactions between variables 1 and 2 (AUTHORN and FACGROUP) and between factors 2 and 3 (FACGROUP and LASTCODE) were highly significant for both the marginal and partial associations. The interaction between factors 1 and 3 (i.e. AUTHORN and LASTCODE) was only significant when the marginal association between the two factors was tested, i.e. when there were no other two-way interactions in the model; the non-significant test of partial associations showed that exclusion of the AUTHORN*LASTCODE interaction from a model containing all two-way interactions did not result in a significant deterioration of the fit of the model.

Exhibit 8.12 : Results of Fitting all K-Factor Interactions (comfu92.sta)

	Degrs.of Freedom	Max.Lik. Chi-squ.	Probab. p	Pearson Chi-squ.	Probab. p
1 (AUTHORN)	6	279.0295	0.000000	392.0980	0.000000
2 (FACGROUP)	11	58.5977	.000000	124.4270	.000000
3 (LASTCODE)	6	6.8257	.337294	7.2750	.296196

Exhibit 8.13 : Tests of Marginal and Partial Association (comfu92.sta)

	Degrs.of Freedom	Prt.Ass. Chi-sqr.	Prt.Ass. p	Mrg.Ass. Chi-sqr.	Mrg.Ass. p
1 (AUTHORN)	2	150.4090	.000000	150.4090	.000000
2 (FACGROUP)	3	75.6200	.000000	75.6200	.000000
3 (LASTCODE)	1	53.0004	.000000	53.0004	.000000
12 (AUTHORN*FACGROUP)	6	20.6797	.002099	27.6295	.000111
13 (AUTHORN*LASTCODE)	2	2.7461	.253351	9.6958	.007851
23 (FACGROUP*LASTCODE)	3	21.2723	.000093	28.2221	.000003

It would appear that the best-fit model should therefore incorporate the 12 (AUTHORN*FACGROUP) and 23 (FACGROUP*LASTCODE) interactions only. Automatic

selection of the best-fit model confirmed the manual selection of significant interactions. A goodness of fit evaluation of the best-fit (12 23) model is presented in Exhibit 5.14 below:

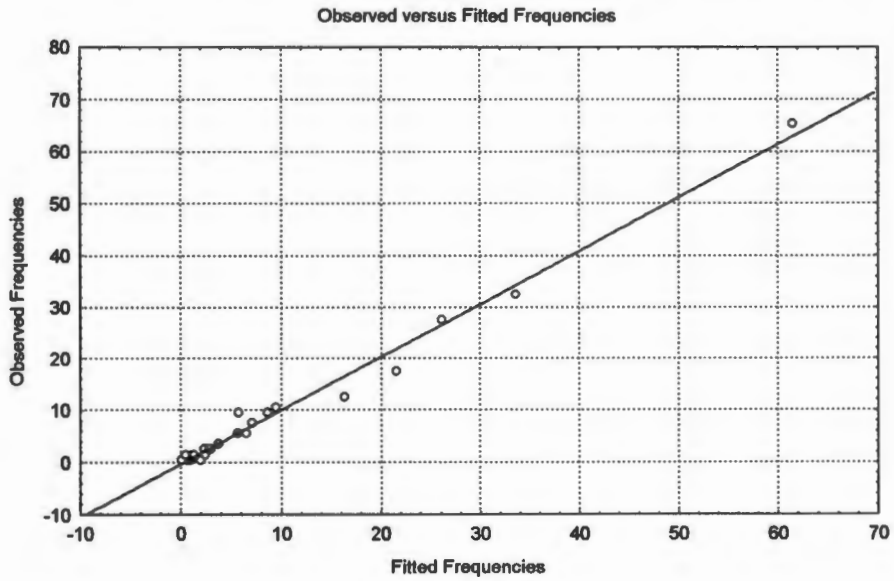
Exhibit 8.14 : Test of goodness of fit of the 12 23 best-fit model

Observed Table: AUTHORN(3) * FACGROUP(4) * LASTCODE(2) (comfu92.sta)
Model: 21,32

	Chi-sqr	df	p
Max Likelihood Chi-square	9.57181	8	.296415
Pearson Chi-square	11.64159	8	.167976

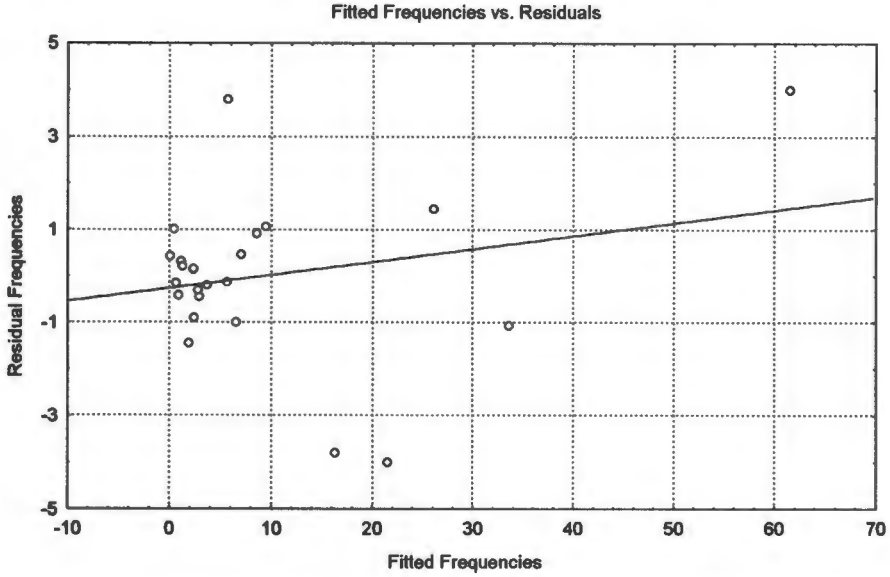
Examination of a plot of the observed frequencies in the three-way table against those fitted under the best-fit model (Exhibit 8.15) implied a close linear relationship with no major outliers:

Exhibit 8.15 : Plot of observed frequencies vs. fitted frequencies under best-fit model



Furthermore, a plot of the fitted frequencies and the residual frequencies under the best-fit model (see Exhibit 8.16) yielded a random scatter, thus lending support to the best-fit (12 23) model :

Exhibit 8.16 : Plot of fitted frequencies vs. residuals under the best-fit model



Prior to interpretation of the effects in the best-fit model, a hierarchical test was carried out to assess the effects of removing the 12 interaction (i.e. the interaction between the design variables AUTHORN and FACGROUP) on the overall model fit. Evaluation of the goodness of fit of the resultant 1 32 model (the main effect of the AUTHORN factor must be specified if the model to be tested does not include this effect - see Exhibit 8.17) showed that there was a significant difference between the observed frequencies and those fitted under the model selected:

Exhibit 8.17 : Test of goodness of fit of a model excluding the interaction between design variables

Observed Table: AUTHORN(3) * FACGROUP(4) * LASTCODE(2) (comfu92.sta)
Model: 1,32

	Chi-sqr	df	p
Max Likelihood Chi-square	37.20134	14	.000691
Pearson Chi-square	54.38196	14	.000001

A second hierarchical test was carried out in order to assess the improvement in fit achieved by incorporating the 13 (AUTHORN*LASTCODE) interaction which was not incorporated in the automatic and manual best-fit model. A test of the goodness of fit of the 12 23 13 model is presented in Exhibit 8.18 below:

Exhibit 8.19 : Test of goodness if fit of a model incorporating all two-way interactions

Observed Table: AUTHORN(3) * FACGROUP(4) * LASTCODE(2) (comfu92.sta)
Model: 21,32,31

	Chi-sqr	df	p
Max Likelihood Chi-square	6.821433	6	.337706
Pearson Chi-square	7.355468	6	.289254

Computation of the chi-square difference between the best-fit model and the model including all two-way interactions (chi-square difference = 2.750377 at df difference = 2) showed that the 13 association did not significantly improve the model fit, and hence that the best-fit model adequately described the relationships in the three-way cross-tabulation examined in this analysis.

Interpretation of the model was carried out by examining the marginal tables pertaining to the best-fit model. Whilst the marginal table between the design variables AUTHORN and FACGROUP (Exhibit 8.20) indicated that there were relatively more low faculty points DET students than in the other 2 matric authority groups, the nature of the effects depicted in the LASTCODE*FACGROUP (Exhibit 8.21) are of particular interest:

Exhibit 8.20 : Marginal table of AUTHORN*FACGROUP

Marg. Tabl. (freq+delta):AUTHORN by FACGROUP (comfu92.sta)

	AUTHORN DET	AUTHORN CO	AUTHORN OMA	Total
1	10.00000	4.00000	8.0000	22.0000
2	5.00000	22.00000	83.0000	110.0000
3	3.00000	11.00000	43.0000	57.0000
4	1.00000	7.00000	28.0000	36.0000
Total	19.00000	44.00000	162.0000	225.0000

Exhibit 8.21 : Marginal table of FACGROUP*LASTCODE

Marg. Tabl. (freq+delta):FACGROUP by LASTCODE (comfu92.sta)

	FACGROUP 1	FACGROUP 2	FACGROUP 3	FACGROUP 4	Total
QUA	6.50000	81.5000	44.50000	33.50000	166.0000
REN	15.50000	28.5000	12.50000	2.50000	59.0000
Total	22.00000	110.0000	57.00000	36.00000	225.0000

Calculation of the QUA:REN ratios for each faculty grouping indicated a steady increase (from 0.4 for FACGROUP=1 to 13.4 for FACGROUP=4). The overall QUA:REN rate calculated from the marginal table (2.81) was similar to that for FACGROUP=2 and was thus severely affected by the relatively low graduation rates and high rates of academic exclusion amongst the lower faculty point groupings.

FACULTY OF SCIENCE (BSc cohort)

Chi-square analysis was initially employed in order to establish whether or not there were statistically significant associations between either of the design variables (AUTHORN and FACGROUP) and the two-level response variable LASTCODE within the BSc cohort. The results of these analyses (see Exhibit 8.22) reflect highly significant two-way associations between each of the design variables and the response variable.

Exhibit 8.22 : Summary of chi-square analyses

Design variable	Pearson Chi-square	df	p	Significant?
AUTHORN	36.395	df=2	p=.00000	YES
FACGROUP	33.282	df=4	p=.00000	YES

The results of fitting all k-factor interactions within the three-way cross-tabulation (Exhibit 8.23) showed that there was a significant improvement in the fit of the model when all two-way interactions were included ($p<.05$), but that the improvement in adding the three-way interaction (i.e. AUTHORN*FACGROUP*LASTCODE) was not significant ($p>.05$). At this point, it would appear that the best-fit model would contain some two-way interactions, but not the three-way interaction.

Exhibit 8.23 : Results of Fitting all K-Factor Interactions (sci92fu.sta)

	Degrs.of Freedom	Max.Lik. Chi-squ.	Probab. p	Pearson Chi-squ	Probab. p
1 (AUTHORN)	7	90.9919	.000000	105.8193	.000000
2 (FACGROUP)	14	113.0192	.000000	137.4151	.000000
3 (LASTCODE)	8	4.9130	.766829	4.7534	.783573

Tests of marginal and partial association (Exhibit 8.24) showed that all three two-way interactions were highly significant in both tests. The manually fit model would therefore include the AUTHORN*FACGROUP, AUTHORN*LASTCODE and FACGROUP*LASTCODE interactions, i.e. a "12 13 23" model.

Exhibit 8.24 : Tests of Marginal and Partial Association (sci92fu.sta)

	Degrs.of Freedom	Prt.Ass. Chi-sqr.	Prt.Ass. p	Mrg.Ass. Chi-sqr.	Mrg.Ass. p
1 (AUTHORN)	2	63.68015	.000000	63.68015	.000000
2 (FACGROUP)	4	5.13356	.273891	5.13356	.273891
3 (LASTCODE)	1	22.17829	.000002	22.17829	.000002
12 (AUTHORN*FACGROUP)	8	44.73513	.000000	58.62501	.000000
13 (AUTHORN*LASTCODE)	2	19.76463	.000051	33.65445	.000000
23 (FACGROUP*LASTCODE)	4	20.73976	.000358	34.62967	.000001

Automatic selection of the best-fit model confirmed that the least complex model that would fit the three-way observed cross-tabulation involved all two-way interactions. A test of goodness of fit of the best-fit model (see Exhibit 8.25) established that there was no significant difference between the fitted frequencies (under the model) and the observed frequencies in the data ($p>.05$).

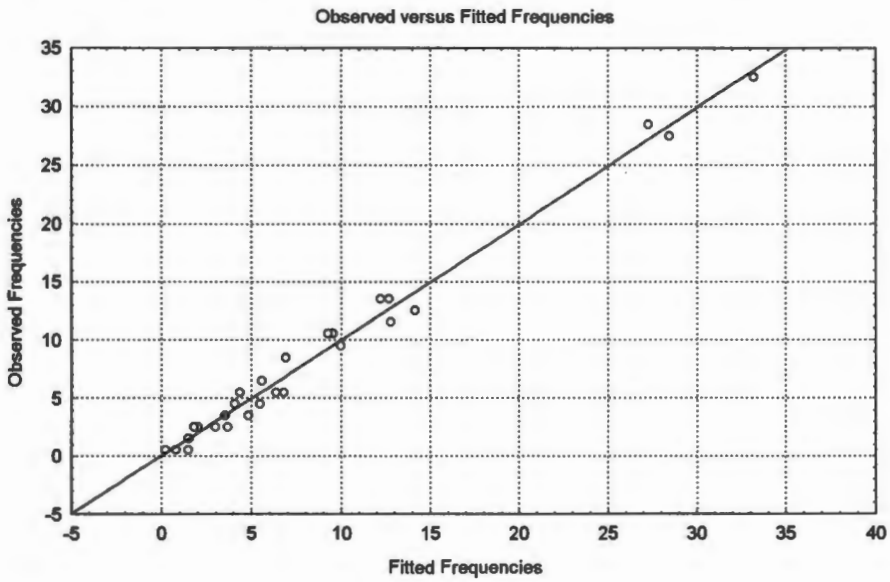
Exhibit 8.25 : Test of goodness of fit of the best-fit model

Observed Table: AUTHORN(3) * FACGROUP(5) * LASTCODE(2) (sci92fu.sta)
Model: 21,31,32

	Chi-sqr	df	p
Max Likelihood Chi-square	4.912055	8	.766924
Pearson Chi-square	4.756423	8	.783255

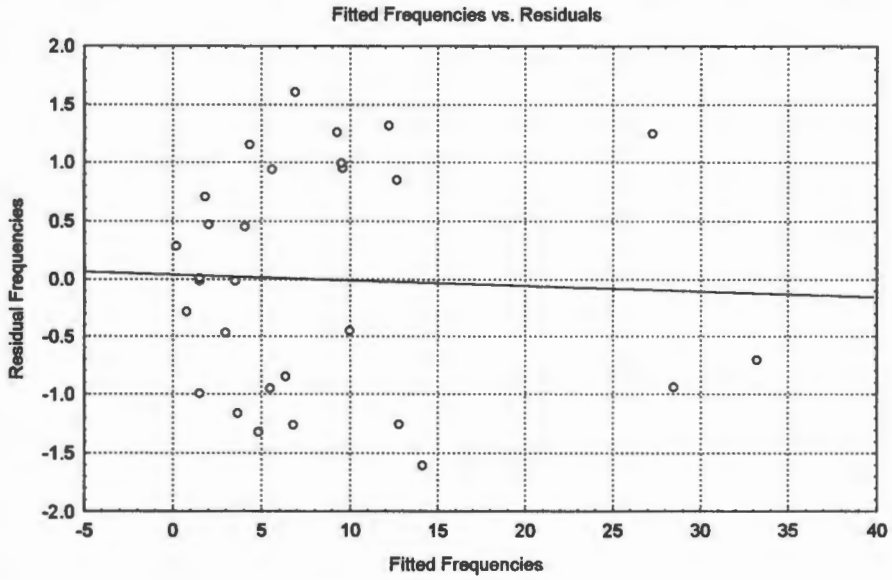
Inspection of a plot of the observed frequencies in the three-way table against the fitted frequencies under the best-fit model (Exhibit 8.26 below) revealed a close linear fit with no major outliers.

Exhibit 8.26 : Plot of observed vs. fitted frequencies under best-fit model



Moreover, a plot of the fitted frequencies against the residual frequencies under the best-fit model (Exhibit 8.27) yielded a scatter of values; this is a further indication of the appropriateness of the best-fit model for the observed three-way table.

Exhibit 8.27 : Plot fitted vs. residual frequencies under best-fit model



Prior to interpretation of the best-fit model, a hierarchical test was performed in order to evaluate the effects of removing the interaction between the design variables (i.e. the 12 or AUTHORN*FACGROUP interaction) from the model. A test of goodness of the resultant 31 32 model (shown in Exhibit 5.28 below) indicated that there was a highly significant difference between the fitted frequencies under the model and those of the observed table. In other words, omission of the interaction between the design variables resulted in an overall lack of fit, indicating that this interaction was essential within the best-fit model.

Exhibit 8.28 : Test of goodness of fit of a model excluding the 12 interaction between design variables

Observed Table: AUTHORN(3) * FACGROUP(5) * LASTCODE(2) (sci92fu.sta)
Model: 31,32

	Chi-sqr	df	p
Max Likelihood Chi-square	49.64809	16	.000026
Pearson Chi-square	54.95010	16	.000004

Interpretation of the best-fit model was performed via inspection of the marginal tables pertaining to the significant interactions (Exhibits 5.29 to 5.30). Exhibit 8.29 demonstrates that there were significantly higher proportions of DET students within the two lower faculty points groups (21/47=0.446 and 17/47=0.362) than within the whole cohort (43/244=0.176 and 41/244=0.168 respectively). The faculty points group distribution amongst the CO students was more favourable with proportions of 8/55=0.145 and 5/55=0.091 in faculty points groups 1 and 2 respectively, but not as favourable as that within the OMA group (where the proportions within faculty points groups 1 and 2 were only 14/142=0,099 and 19/142=0.134 respectively).

Exhibit 8.29 : Marginal table of AUTHORN*FACGROUP

Marg. Tabl. (freq+delta):AUTHORN by FACGROUP (sci92fu.sta)

	AUTHORN DET	AUTHORN CO	AUTHORN OMA	Total
1	21.00000	8.00000	14.0000	43.0000
2	17.00000	5.00000	19.0000	41.0000
3	5.00000	16.00000	40.0000	61.0000
4	3.00000	15.00000	34.0000	52.0000
5	1.00000	11.00000	35.0000	47.0000
Total	47.00000	55.00000	142.0000	244.0000

The marginal table between AUTHORN and LASTCODE (Exhibit 8.30) shows that the QUA:REN ratio increased from 0.329 amongst the DET matric authority group to 1.35 within the CO group. The ratio amongst the OMA group (3.656) was considerably higher than that calculated for the cohort as a whole (158.5/85.5=1.854).

Exhibit 8.30 : Marginal table of AUTHORN*LASTCODE

Marg. Tabl. (freq+delta):AUTHORN by LASTCODE (sci92fu.sta)

	AUTHORN DET	AUTHORN CO	AUTHORN OMA	Total
QUA	15.50000	31.50000	111.5000	158.5000
REN	31.50000	23.50000	30.5000	85.5000
Total	47.00000	55.00000	142.0000	244.0000

Inspection of the marginal table between FACGROUP and LASTCODE (Exhibit 8.31) revealed a progressive increase in the QUA:REN ratio with increasing faculty points group. The calculated ratios were 20.5/22.5=.898 for FACGROUP=1, 19.5/21.5=0.907 for FACGROUP=2, 35.5/25.5=1.392 for FACGROUP=3, 39.5/12.5=3.160 for FACGROUP=4 and 43.5/3.5=12.429 for FACGROUP=5. The QUA:REN ratio calculated from marginal table for the cohort as a whole (158.5/85.5=1.854) therefore lay between the ratios for FACGROUPs 2 and 3. The very low rates of graduation coupled with the high rates of academic exclusion within the lower two faculty points groups which (within this particular cohort) represent a combined proportion of only 34% of the marginal total, therefore had a disproportionately large effect (measured in terms of size of the FACGROUPs) on the QUA:REN ratio within the cohort.

Exhibit 8.31 : Marginal table of FACGROUP*LASTCODE

Marg. Tabl. (freq+delta):FACGROUP by LASTCODE (sci92fu.sta)

	FACGROUP 1	FACGROUP 2	FACGROUP 3	FACGROUP 4	FACGROUP 5	Total
QUA	20.50000	19.50000	35.50000	39.50000	43.50000	158.5000
REN	22.50000	21.50000	25.50000	12.50000	3.50000	85.5000
Total	43.00000	41.00000	61.00000	52.00000	47.00000	244.0000

Although there was no significant three-way interaction between the AUTHORN, FACGROUP and LASTCODE variables, both the matric authority and the faculty points group clearly contributed to the undergraduate academic outcomes within this cohort. Given the dramatic effects of faculty points group on the QUA:REN ratio, the particularly poor graduation rate within the DET group could be largely ascribed to the relatively large numbers of DET students within the lower faculty points groups. Similarly, the intermediate

graduation rate within the CO group reflects the intermediate FACGROUP distribution within this matric authority grouping.

FACULTY OF SOCIAL SCIENCE AND HUMANITIES (BA/BSocSc cohort)

Prior to log-linear modelling, chi-square analysis was used to establish whether or not there were significant associations between either of the design variables and the two-level response variable. The summary of the results of the chi-square testing presented in Exhibit 8.32 confirms that there were no significant associations between either of the design variables (AUTHORN and FACGROUP) and the response variable LASTCODE. The interpretation of these findings is that there were no significant differences in graduation and academic exclusion rates amongst either the matric authority or the faculty points groups. One would expect that the log-linear model of the three-way cross-tabulation (AUTHORN*LASTCODE*FACGROUP) would not incorporate interactions between either of the design variables and the response variable LASTCODE.

Exhibit 8.32 : Summary of Chi-square analyses

Design variable	Pearson Chi-square	df	p	Significant?
AUTHORN	5.343	df=2	p=.06916	NO
FACGROUP	2.755	df=4	p=.59958	NO

Within Statistica's log-linear module, an examination of all k-factor interactions (see Exhibit 8.33) showed that there was a significant effect in including all two-way interactions, but that the improvement in model fit when adding the three-way interaction is not significant.

Exhibit 8.33 : Results of Fitting all K-Factor Interactions (sshfu92.sta)

	Degrs.of Freedom	Max.Lik. Chi-squ.	Probab. p	Pearson Chi-squ	Probab. p
1 (AUTHORN)	7	434.7678	0.000000	740.4845	0.000000
2 (FACGROUP)	14	92.6376	.000000	100.8886	.000000
3 (LASTCODE)	8	4.6098	.798338	5.8187	.667531

Inspection of the tests of marginal and partial association (Exhibit 8.34) indicated that the only significant two-way interaction was that between factors 1 and 2 (corresponding to the design variables AUTHORN and FACGROUP). This interaction was highly significant in both the test of partial association and that of marginal association (p=.0000 in both cases). At this point, a manual model containing the interaction between the design variables and the main effect of the LASTCODE variable could be tested.

Exhibit 8.34 : Tests of marginal and partial association

Tests of Marginal and Partial Association (sshfu92.sta)

	Degrs.of Freedom	Prt.Ass. Chi-sqr.	Prt.Ass. p	Mrg.Ass. Chi-sqr.	Mrg.Ass. p
1 (AUTHORN)	2	99.3790	.000000	99.3790	.000000
2 (FACGROUP)	4	160.2536	.000000	160.2536	.000000
3 (LASTCODE)	1	175.1350	.000000	175.1350	.000000
12 (AUTHORN*FACGROUP)	8	85.8494	.000000	85.0047	.000000
13 (AUTHORN*LASTCODE)	2	4.9874	.082619	4.1427	.126032
23 (FACGROUP*LASTCODE)	4	3.4903	.479368	2.6456	.618777

Automatic selection of the best-fit model confirmed that the 12, 3 (i.e. the AUTHORN*FACGROUP, LASTCODE) model provided a satisfactory fit with regard to the observed table. A test of goodness of fit of the selected best-fit model (Exhibit 8.35) established that there was no significant difference (p>.05) between the frequencies under the model and those within the observed three-way table.

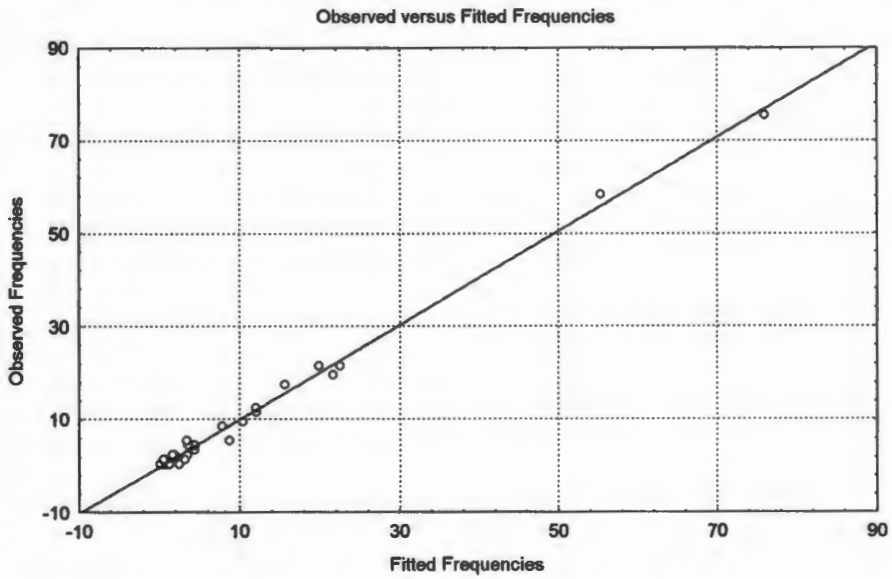
Exhibit 8.35 : Test of goodness of fit of the best-fit model

Observed Table: AUTHORN(3) * FACGROUP(5) * LASTCODE(2) (sshfu92.sta)
Model: 21,3

	Chi-sqr	df	p
Max Likelihood Chi-square	12.24274	14	.586819
Pearson Chi-square	12.80985	14	.541561

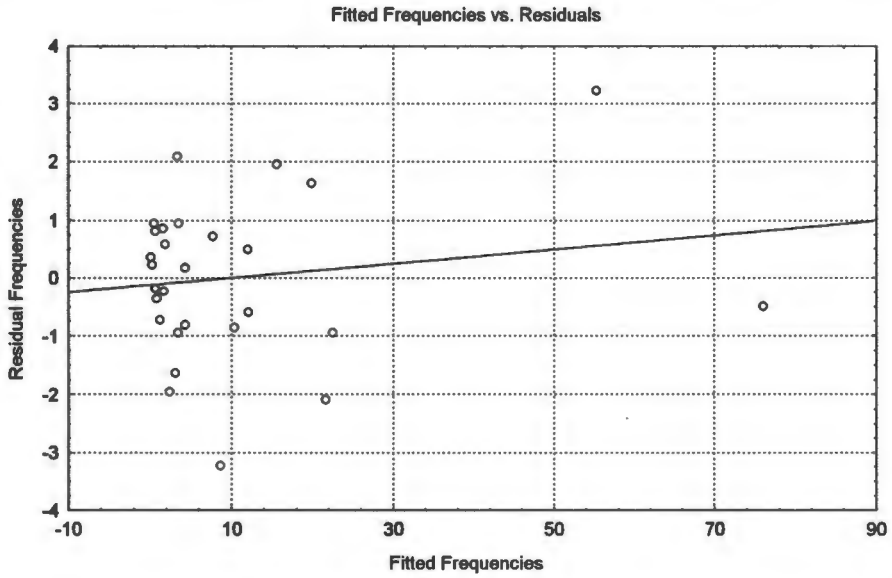
An examination of a plot of the observed frequencies against the fitted frequencies under the best-fit model (Exhibit 8.36) showed a close linear fit, and that there were no major outliers:

Exhibit 8.36 : Plot of observed vs. fitted frequencies under best-fit model



A plot of the residual frequencies against the fitted frequencies under the best-fit model (see Exhibit 8.37) produced a random scatter, adding further support for the best-fit model.

Exhibit 8.37 : Plot of fitted vs. residual frequencies under the best-fit model



Before interpreting the results of the model, two hierarchical tests were performed (i) to test for a significant improvement in including the AUTHORN*LASTCODE (13) interaction and (ii) to test for a significant improvement in the model fit by including the FACGROUP*LASTCODE (23) interaction. A test of the goodness of fit of the 12 13 model (presented in Exhibit 8.38 below) showed that the chi-square difference between the best-fit model and the 21, 31 model (12.2427-8.100108= 4.142592 at df difference=2, p>.05) was not significant, indicating that there was no significant difference between the best-fit model and a model incorporating the 13 (AUTHORN*LASTCODE) interaction.

Exhibit 8.38 : Test of goodness of fit of a model including the AUTHORN*LASTCODE interaction

Observed Table: AUTHORN(3) * FACGROUP(5) * LASTCODE(2) (sshfu92.sta)
Model: 21,31

	Chi-sqr	df	p
Max Likelihood Chi-square	8.100108	12	.777240
Pearson Chi-square	8.350852	12	.757129

A test of goodness of fit of the 21 32 model (i.e. including the FACGROUP*LASTCODE interaction) is presented in Exhibit 8.39. Computation of the chi-square difference between the best-fit model and the 21 32 model (chi-square difference=12.2427-9.59723= 2.64547 at df difference=12-10=2, p>.05) showed that there was no significant difference in the fit of the best-fit and the 21 32 model.

Exhibit 8.39 : Test of goodness of fit of a model including the FACGROUP*LASTCODE interaction

Observed Table: AUTHORN(3) * FACGROUP(5) * LASTCODE(2) (sshfu92.sta)
Model: 21,32

	Chi-sqr	df	p
Max Likelihood Chi-square	9.59723	10	.476532
Pearson Chi-square	11.00403	10	.357244

It would therefore appear that, in terms of the best-fit model, the variable LASTCODE (representing the QUA and REN levels within the range of final undergraduate academic outcomes) was independent of both the matric authority and faculty points grouping variables. In other words, neither the matric authority group (AUTHORN) nor the faculty points group (FACGROUP) contributed significantly to the rates of graduation and academic exclusion observed within the BA/BSocSc cohort. The interaction between the AUTHORN and FACGROUP design variables (depicted in Exhibit 8.40 below) showed that the proportion of FACGROUP=1 and FACGROUP=2 students within the DET group was particularly high (0.548) in comparison with the equivalent proportions within the CO and OMA groups (0.068 and 0.056 respectively) and within the cohort as a whole (0.197). The prevalence of low faculty points students within the DET matric authority group did not have an adverse effect on the overall rate of graduation within this group, however.

Exhibit 8.40 : Marginal table of AUTHORN*FACGROUP

Marg. Tabl. (freq+delta):AUTHORN by FACGROUP (sshfu92.sta)

	AUTHORN DET	AUTHORN CO	AUTHORN OMA	Total
1	14.00000	1.00000	5.0000	20.0000
2	26.00000	2.00000	5.0000	33.0000
3	23.00000	25.00000	88.0000	136.0000
4	9.00000	12.00000	64.0000	85.0000
5	1.00000	4.00000	18.0000	23.0000
Total	**73.00000**	**44.00000**	**180.0000**	**297.0000**

APPENDIX 2

THE DEVELOPMENT OF MULTIPLE DISCRIMINANT MODELS AND THE IDENTIFICATION OF PREDICTOR VARIABLES OF UNDERGRADUATE ACADEMIC OUTCOMES.

The development of the multiple discriminant models for each of the selected cohorts is discussed below. The method used in the multiple discriminant analyses is described in Chapter 4 of this report. In all of the analyses, the independent variables selected were the SEXN, AUTDUM1, AUTDUM2, FACPOINT, ENGPTS and MATPTS variables; the contents of these variables and their coding are described in Chapter 4 of this report. The grouping variable in all of the analyses was the final undergraduate outcome variable LASTCODE. Only the "QUA" and "REN" codes within the variable LASTCODE were included in the analyses, which were carried out in order to discriminate between the graduates and the academic exclusions within each of the cohorts.

FACULTY OF ARTS (BA cohort)

In preparation for multiple discriminant analysis, the correlations between the variables in the analysis (i.e. SEXN, AUTDUM1, AUTDUM2, FACPOINT, ENGPTS and MATPTS) were examined. Exhibit 9.1 below shows that although the correlations were generally weak, there was a moderately strong correlation between the variables FACPOINT and MATPTS. Because the variable FACPOINT was a key variable in this analysis, the correlated variable MATPTS (which is the Swedish points equivalent of the matric Mathematics symbol) was subsequently excluded from the multiple discriminant analysis.

Exhibit 9.1 : Coefficients of linear correlation between independent variables

Correlations (Total) (art92fu.sta)

	SEXN	AUTDUM1	AUTDUM2	FACPOINT	ENGPTS	MATPTS
SEXN	1.00	-.19	-.07	.11	-.02	.27
AUTDUM1	-.19	1.00	-.13	-.21	-.20	-.13
AUTDUM2	-.07	-.13	1.00	-.43	-.09	-.56
FACPOINT	.11	-.21	-.43	1.00	.43	.73
ENGPTS	-.02	-.20	-.09	.43	1.00	.11
MATPTS	.27	-.13	-.56	.73	.11	1.00

Prior to the analysis, a range of descriptive statistics pertaining to the variables in the analysis was examined. A tabulation of the means and number of valid cases (see Exhibit 9.2 below) shows that the group means (i.e. for the "QUA" and "REN" groups within the variable LASTCODE) for the variables SEXN and ENGPTS were very similar, but that there was separation between the group means for the remainder of the variables.

Exhibit 9.2 : Group means, overall means and standard deviations of independent variables

Means (art92fu.sta)

	SEXN	AUTDUM1	AUTDUM2	FACPOINT	ENGPTS	Valid N
QUA	.452381	.119048	.190476	34.26984	6.492064	126
REN	.520000	.320000	.280000	30.24000	6.080000	25
All Grps	.463576	.152318	.205298	33.60265	6.423841	151

Standard Deviations (art92fu.sta)

	SEXN	AUTDUM1	AUTDUM2	FACPOINT	ENGPTS	Valid N
QUA	.499714	.325137	.394244	5.883418	1.001966	126
REN	.509902	.476095	.458258	4.867238	.909212	25
All Grps	.500331	.360525	.405263	5.907147	.996240	151

In order to check that the underlying assumption of normality was not violated for any of the variables in the analysis, a categorised normal probability plot was produced for each of the independent variables in the analysis (Exhibit 9.3a - e). Inspection of these plots indicated that the variables in the analysis were basically normally distributed and that there were no major outliers (with the possible exception of within the variable ENGPTS, Exhibit 9.3e) which could have a large impact on the means, possibly resulting in a high level of correlation between the means and their standard deviations.

Exhibit 9.3a : Normal probability plot of the variable SEXN

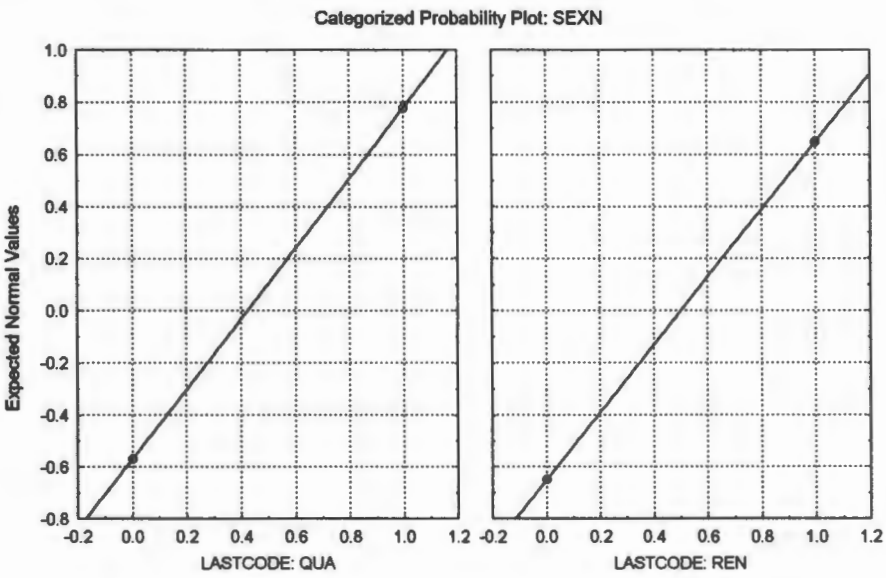


Exhibit 9.3b : Normal probability plot of the variable AUTDUM1

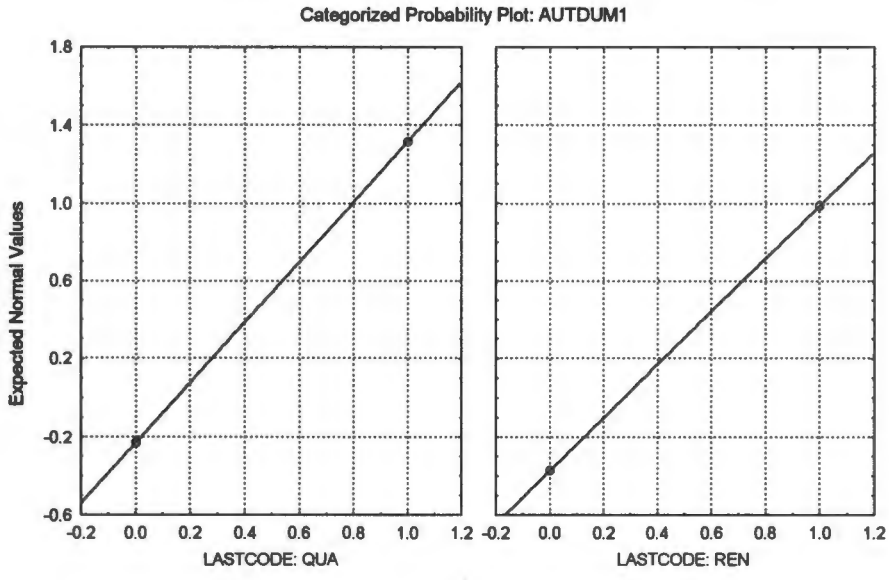


Exhibit 9.3c : Normal probability plot of the variable AUTDUM2

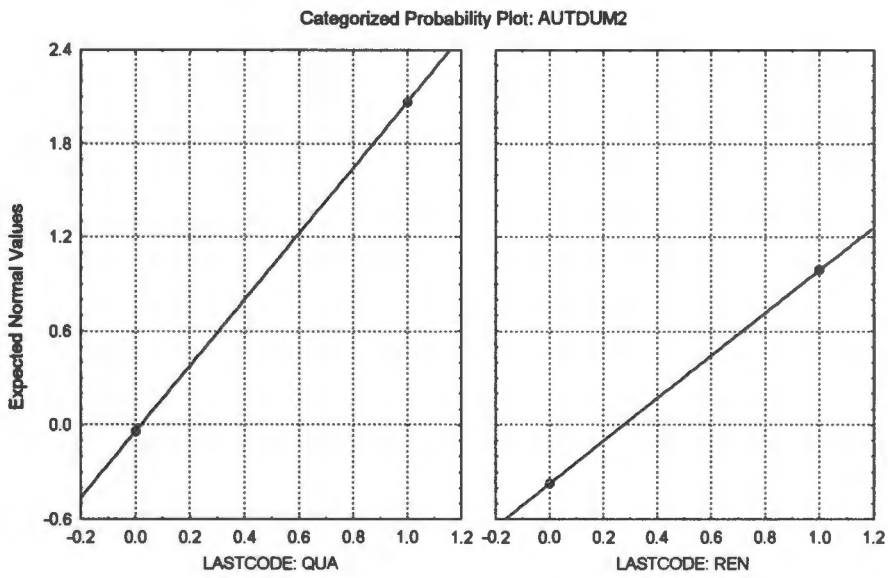


Exhibit 9.3d : Normal probability plot of the variable AUTDUM2

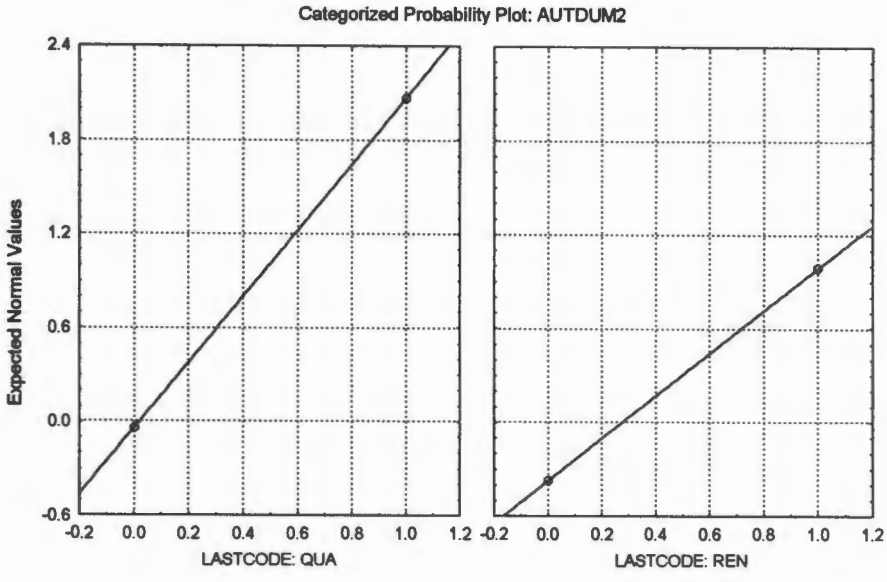


Exhibit 9.3e: Normal probability plot of the variable FACPOINT

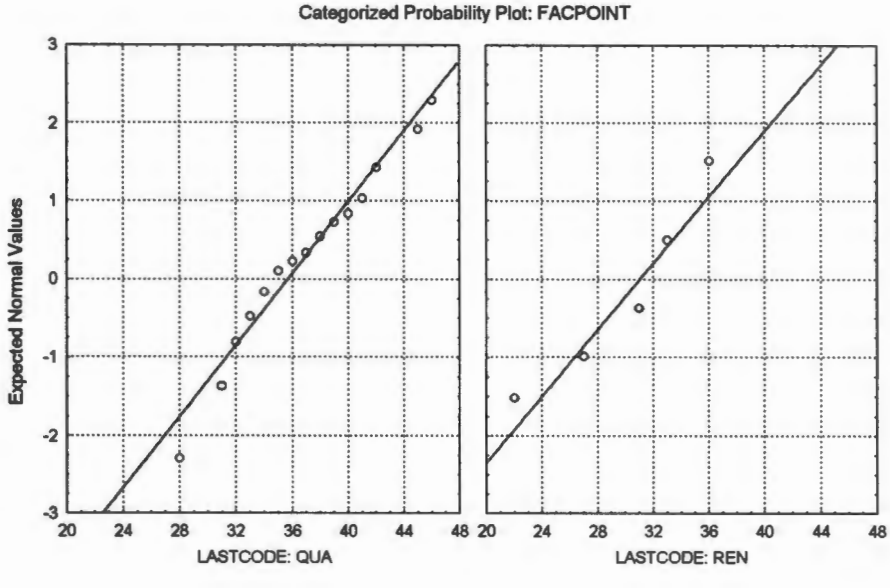
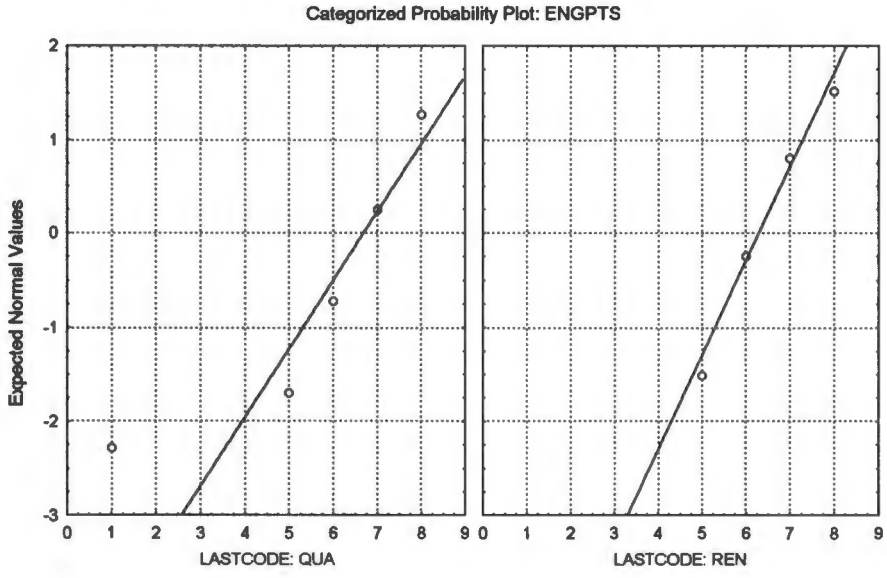


Exhibit 9.3e : Normal probability plot of the variable ENGPTS



Categorised box-and-whisker plots of the variables in the analysis are presented in Exhibits 9.4a - e below. Examination of the box-and-whisker plots indicates that the data within the variables tended to be widely dispersed around the group means, and that the FACPOINT variable and the two matriculation authority dummy variables (AUTDUM1 and AUTDUM2) appeared to achieve the best separation of the two groups within the dependent variable LASTCODE. These three variables would therefore seem to be the most likely to be included in the multiple discriminant model.

Exhibit 6.4 a: Box-and-whisker plot of the variable SEXN by group

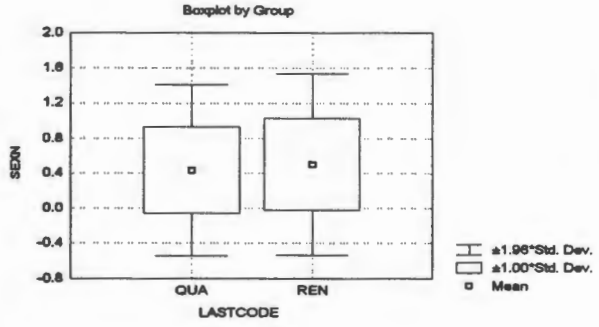


Exhibit 6.4 b: Box-and-whisker plot of the variable AUTDUM1 by group

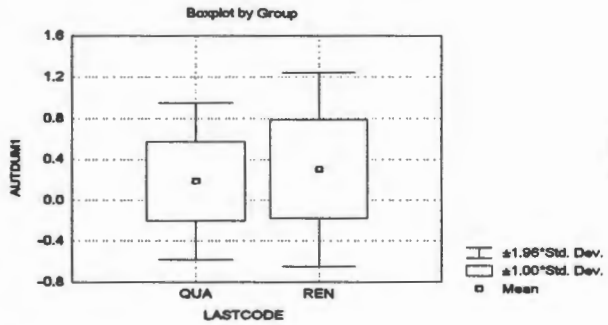


Exhibit 6.4 c: Box-and-whisker plot of the variable AUTDUM2 by group

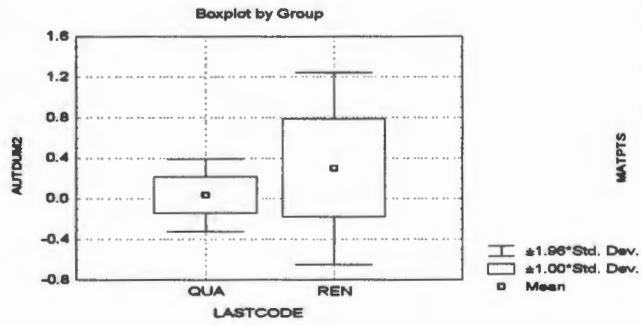


Exhibit 6.4d : Box-and-whisker plot of the variable FACPOINT by group

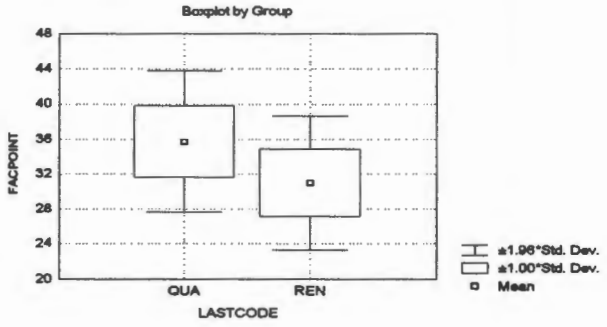


Exhibit 6.4e : Box-and-whisker plot of the variable ENGPTS by group

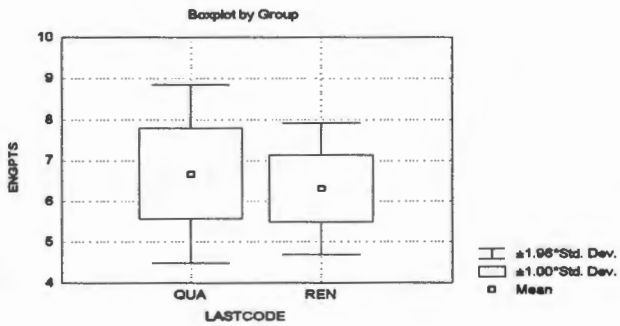
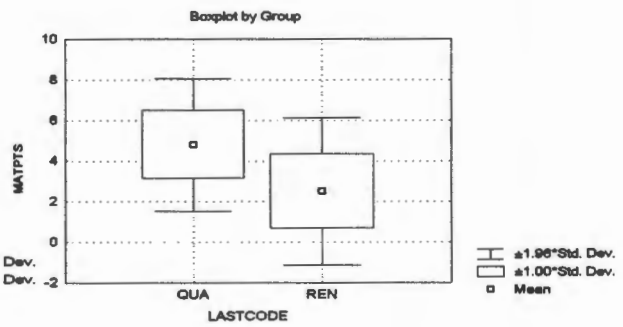


Exhibit 6.4f : Box-and-whisker plot of the variable MATPTS by group



A summary of the results of the forward stepwise analysis (with F to enter set at 3 and the tolerance at the default setting of 0.01) are shown in Exhibit 9.5 below. In terms of the multiple discriminant model, the variables FACPOINT and AUTDUM1 (where students within the CO matriculation authority were coded as 1) were identified as the best predictors of "QUA" or "REN" group membership within the grouping variable LASTCODE. The standard statistic used in the evaluation of the discriminatory power (Wilks' *lambda*, which varies from 0.0 for perfect discriminatory power to 1.0 for no discriminatory power (StatSoft, 1995: 3084)) for the present model was 0.89554 (see Exhibit 9.5), indicating weak discriminatory power. Examination of the partial *lambda's* shows that the variable FACPOINT (with the lower partial lambda) provided the greater discriminatory power within the model.

Exhibit 9.5 : Summary of multiple discriminant analysis

Discriminant Function Analysis Summary (art92fu.sta)
Step 2, N of vars in model: 2; Grouping: LASTCODE (2 grps)
Wilks' Lambda: .89554 approx. F (2,148)=8.6319 p< .0003

	Wilks' Lambda	Partial Lambda	F-remove (1,148)	p-level	Toler.	1-Toler. (R-Sqr.)
FACPOINT	.956793	.935979	10.12319	.001784	.999639	.000361
AUTDUM1	.931414	.961482	5.92903	.016083	.999639	.000361

However, examination of the distances between the "QUA" and "REN" groups in terms of the model (see Exhibit 9.6 below) indicates a significant discrimination between the two groups at the 95% confidence level.

Exhibit 9.6 : Evaluation of distances between group means

F-values; df = 2,148 (art92fu.sta)

	QUA	REN
QUA		8.443039
REN	8.443039	

p-levels (art92fu.sta)

	QUA	REN
QUA		.000337
REN	.000337	

Having identified the variables FACPOINT and AUTDUM1 (where the CO matric authority has been coded as 1) as the best discriminators between the "QUA" and the "REN" groups within the dependent variable LASTCODE, and having assessed the discriminatory power of the model, the model's potential in terms of classification accuracy was assessed. A classification matrix, showing the percentage correct (and, by deduction, the percentage incorrect) classifications for each group in terms of the model classification functions, is presented in Exhibit 9.7. The classification matrix indicates that 84% of all cases were classified correctly, but that a very much higher level of classification accuracy was achieved within the "QUA" group (99% correct), as opposed to 4% correct within the "REN" group. The

total number of cases in the classification matrix includes all those cases where there were no missing values for either of the variables in the multiple discriminant model. (Note: The *a priori* probabilities used in the generation of the classification matrix were set to be proportional to group size).

Exhibit 9.7 : Classification matrix under multiple discriminant model

Classification Matrix (art92fu.sta)
Rows: Observed classifications
Columns: Predicted classifications

	Percent Correct	QUA p=.83444	REN p=.16556
QUA	99.21875	127	1
REN	4.00000	24	1
Total	83.66013	151	2

The classification functions which determine likely group membership are shown in Exhibit 9.8. These classification functions could be applied to existing or to new data cases; the likely group membership of each case would be determined on the basis of the group for which the highest classification score is computed. The general formula for application of the classification functions (to be computed for each group in the two group classification) is[1]:

$$s_i = c_i + w_{i1} x_1 + w_{i2} x_2$$

where s is the classification score, i denotes the group, c is the constant, w_{i1} is the weighting for the variable FACPOINT and x_1 is the observed value of the variable FACPOINT, w_{i2} is the weighting for the variable AUTDUM1 and x_2 is the observed value of the variable AUTDUM1 in a particular case.

Exhibit 9.8 : Classification functions under multiple discriminant model

Classification Functions; grouping: LASTCODE (art92fu.sta)

	QUA p=.50000	REN p=.50000
FACPOINT	1.1671	1.0351
AUTDUM1	1.2900	2.8567
Constant	-20.7074	-16.8005

However, in view of the very poor classification accuracy with regard to members of the "REN" group, the application of the classification functions to new cases has little potential value. Moreover, the fact that the classification matrix shown in Exhibit 9.7 is a *post hoc* classification (based on the same cases from which the classification functions were calculated) must be borne in mind, and it is likely that if the classification functions were to be

[1] The general form and application of the classification functions is similar in all four of the multiple discriminant models described in Appendix 1. These functions differ only in the nature of the independent variables included in the classification functions and the numerical equivalents of the weighting factors, which are tabulated within the classification functions pertaining to each. The application of the classification functions is therefore not described again in this Appendix.

applied to a new set of cases (such as members of the 1993 FU cohort), even poorer predictive accuracy would be obtained.

An examination of the classification functions in Exhibit 9.8 reveals that cases with high values for the faculty points variable FACPOINT, and values for AUTDUM1=0 would be classified as members of the "QUA" group. Conversely, those cases with a combination of low values for the FACPOINT variable and AUTDUM1=1 (i.e. matriculated within the CO matric authority) would be classified as members of the "REN" group.

FACULTY OF COMMERCE (BCom cohort)

In preparation for multiple discriminate analysis, the correlations between the independent variables in the analysis (i.e. SEXN, AUTDUM1, AUTDUM2, FACPOINT, ENGPTS and MATPTS) were examined (see Exhibit 9.9). Whilst the correlations between the variables generally were weak, it was noted that there was a moderate correlation between the FACPOINT (faculty points) and MATPTS (the Swedish points equivalent of matric Mathematics symbol). Exclusion of the MATPTS variable from the analysis was found not to affect either Wilks' *lambda* for the model or the variables included in the model; both FACPOINT and MATPTS were therefore included in the analysis.

Exhibit 9.9 : Coefficients of linear correlation between independent variables

Correlations (Total) (comfu92.sta)

	SEXN	AUTDUM1	AUTDUM2	FACPOINT	ENGPTS	MATPTS
SEXN	1.00	-.26	.06	-.05	-.23	.05
AUTDUM1	-.26	1.00	-.13	.00	-.12	-.07
AUTDUM2	.06	-.13	1.00	-.32	.15	-.22
FACPOINT	-.05	.00	-.32	1.00	.42	.69
ENGPTS	-.23	-.12	.15	.42	1.00	-.03
MATPTS	.05	-.07	-.22	.69	-.03	1.00

A range of descriptive statistics pertaining to the matric-related variables selected for the model were explored prior to the multiple discriminant analysis. A tabulation of the variable means and standard deviations, and the means for each variable within the "QUA" and "REN groups in the dependent variable LASTCODE, is presented in Exhibit 9.10 below. It was noted that the group means for the variable FACPOINT differed substantially; there were also marked differences in the group means for the matriculation authority dummy variables AUTDUM1 and AUTDUM2.

Exhibit 9.10 : Group means, overall means and standard deviations of independent variables

Means (comfu92.sta)

	SEXN	AUTDUM1	AUTDUM2	FACPOINT	ENGPTS	MATPTS	Valid N
QUA	.630573	.171975	.044586	50.41401	6.191083	6.165605	157
REN	.641509	.245283	.150943	46.24528	5.924528	5.509434	53
All Grps	.633333	.190476	.071429	49.36190	6.123809	6.000000	210

Standard Deviations (comfu92.sta)

	SEXN	AUTDUM1	AUTDUM2	FACPOINT	ENGPTS	MATPTS	Valid N
QUA	.484194	.378566	.207053	4.426095	.801711	1.153852	157
REN	.484146	.434372	.361420	4.340798	.702990	1.218805	53
All Grps	.483046	.393615	.258155	4.754497	.784969	1.202072	210

In order to establish that the assumption of normality was not violated within the analysis, categorised normal probability plots of the independent variables were inspected (Exhibit 9.11a - f). With the possible exception of the variable MATPTS, these normal probability plots showed that the independent variables were basically normally distributed and that

there were no major outliers which could adversely affect the relationships between the means and their standard deviations.

Exhibit 9.11a : Normal probability plot of the variable SEXN

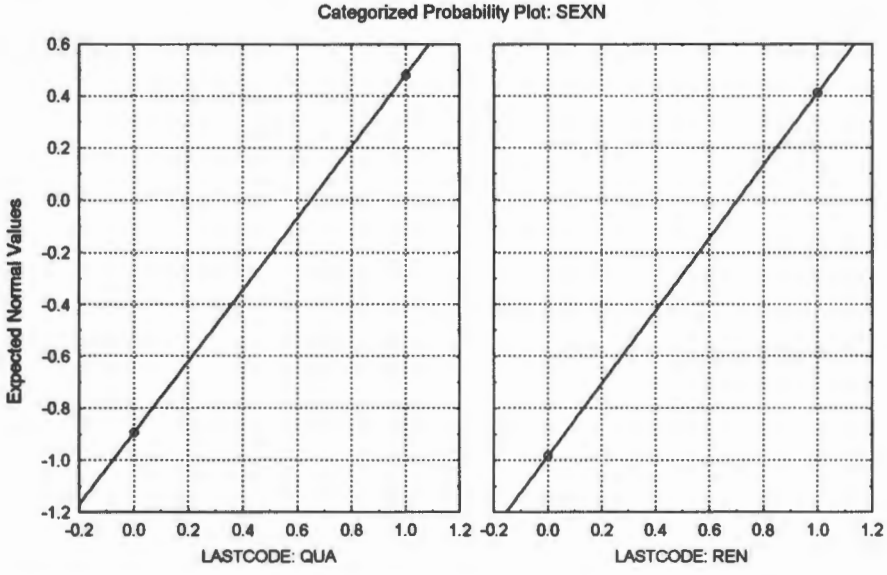


Exhibit 9.11b : Normal probability plot of the variable AUTDUM1

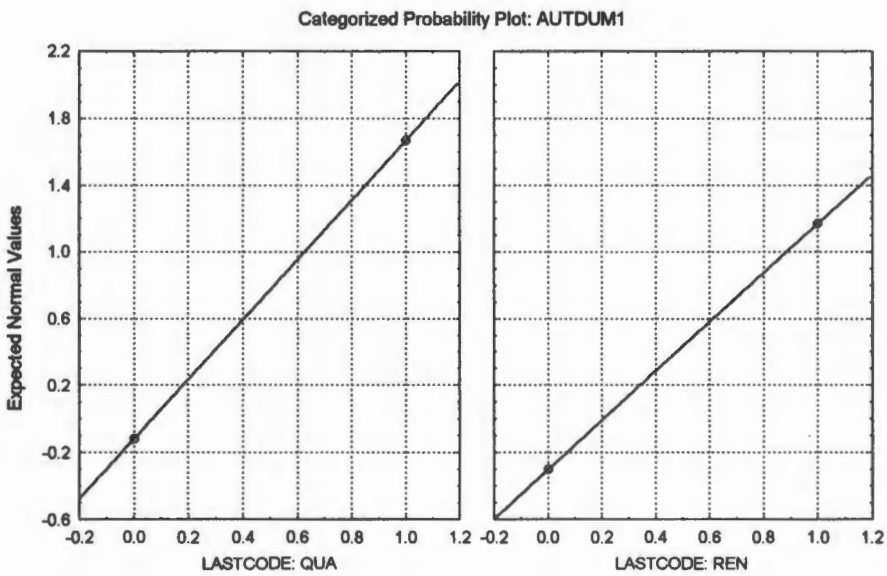


Exhibit 9.11c : Normal probability plot of the variable AUTDUM2

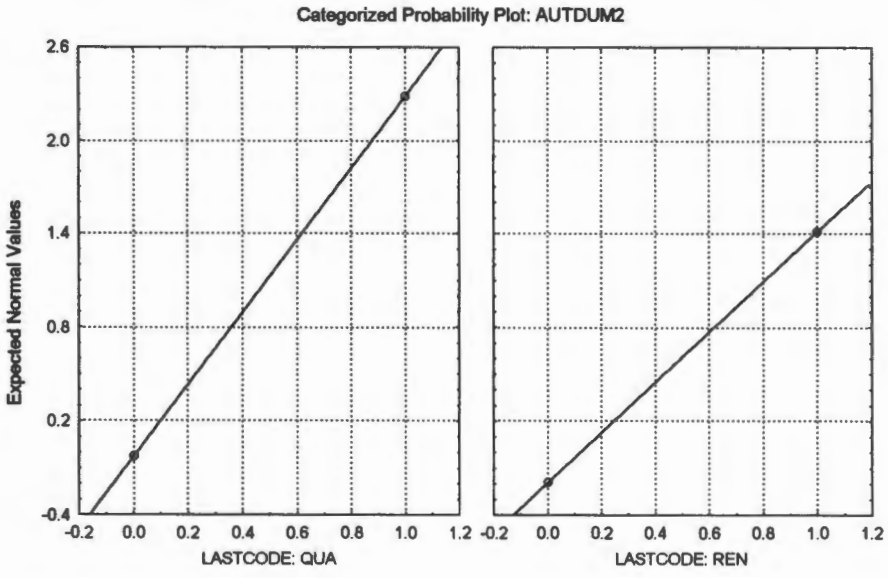


Exhibit 9.11d : Normal probability plot of the variable FACPOINT

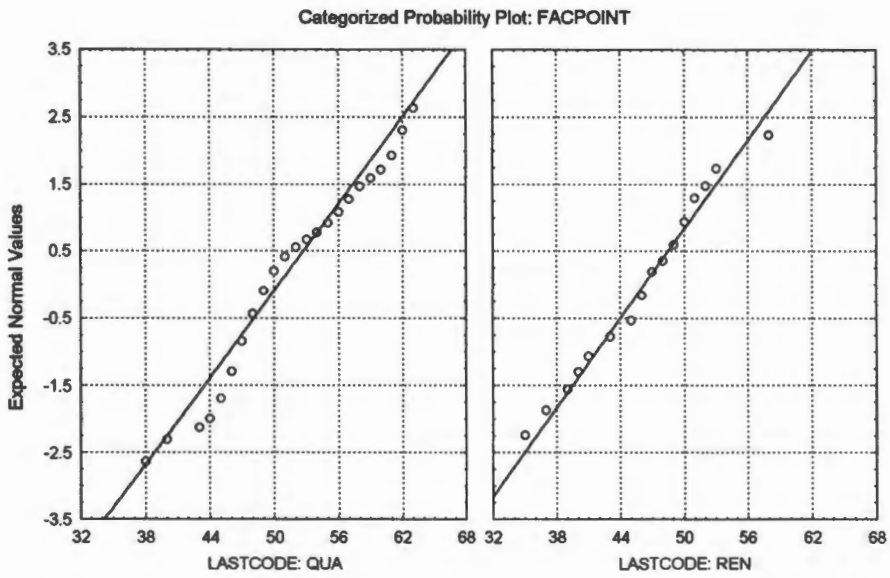


Exhibit 6.12a : Box-and-whisker plot of the variable SEXN by group

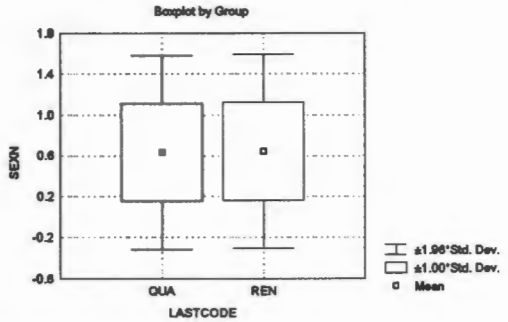


Exhibit 6.12b : Box-and-whisker plot of the variable AUTDUM1 by group

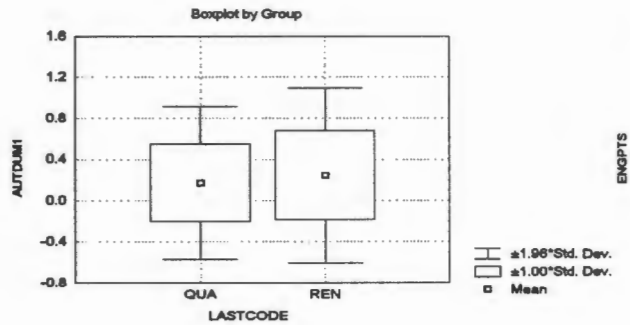


Exhibit 6.12c : Box-and-whisker plot of the variable AUTDUM2 by group

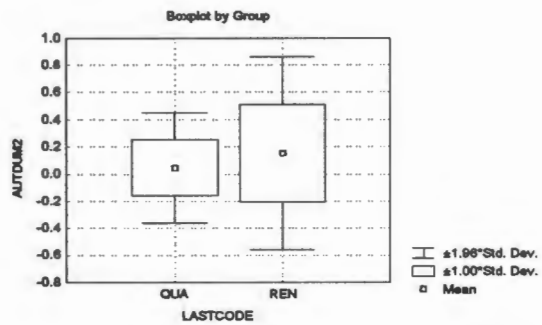


Exhibit 6.12d : Box-and-whisker plot of the variable FACPOINT by group

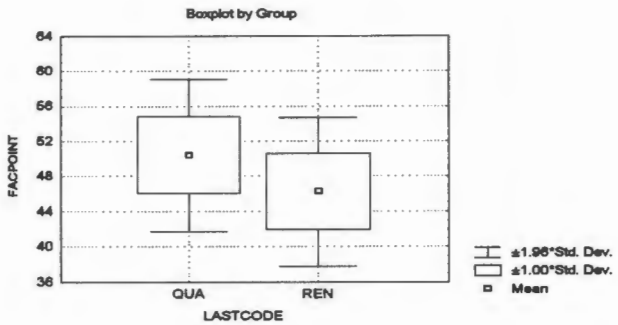


Exhibit 6.12e : Box-and-whisker plot of the variable ENGPTS by group

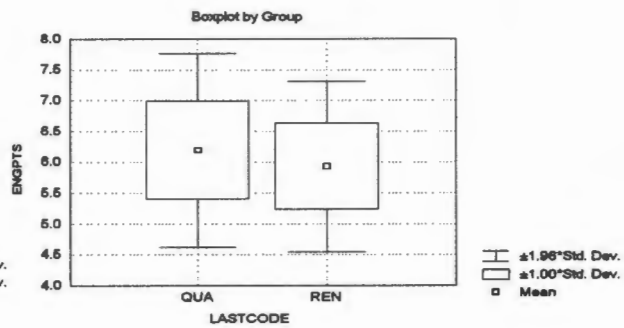


Exhibit 6.12f : Box-and-whisker plot of the variable MATPTS by group

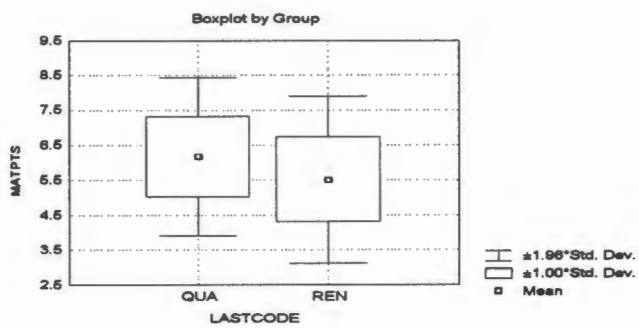


Exhibit 9.11e : Normal probability plot of the variable MATPTS

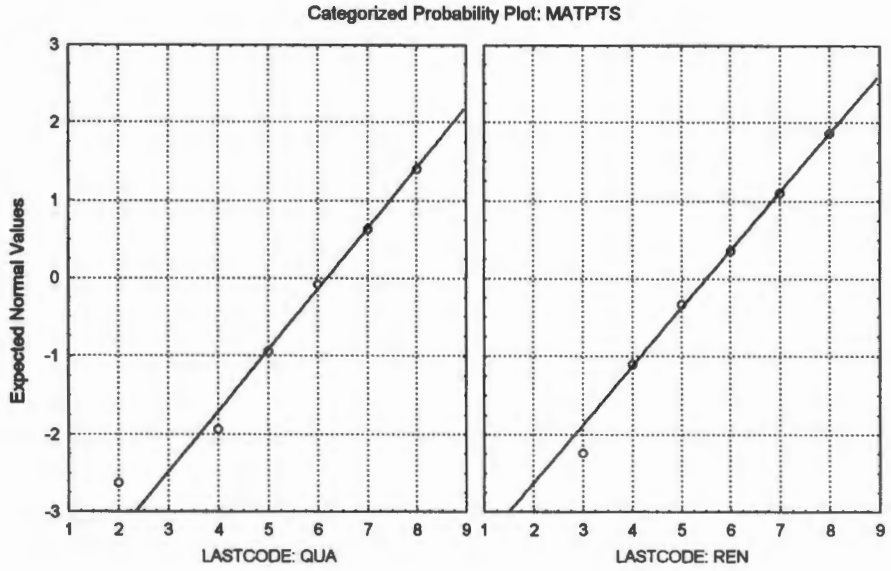
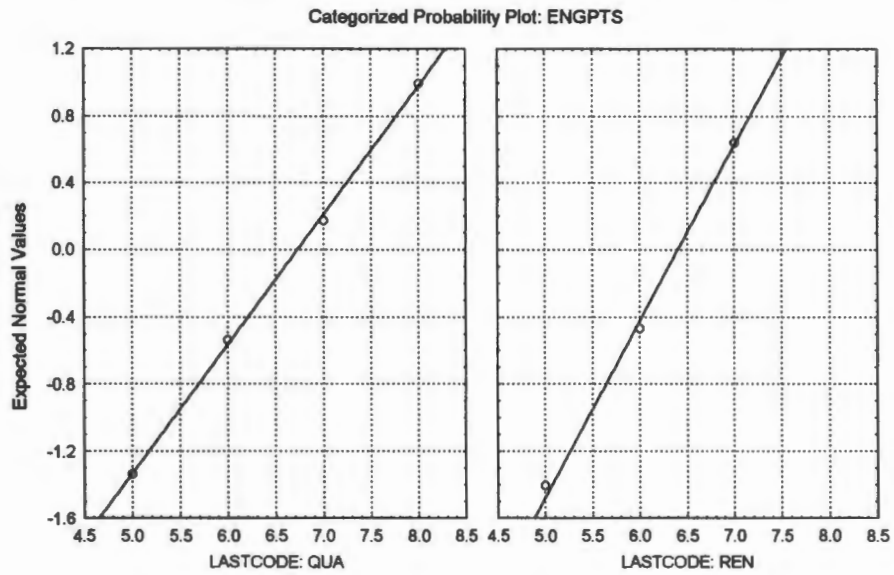


Exhibit 9.11f : Normal probability plot of the variable ENGPTS



Examination of categorised box-and-whisker plots for each of the independent variables in the model (Exhibit 9.12a - f) indicated that the data within the variables tended to be widely dispersed around the group means. There were, however, apparent separations between the group means for the variables FACPOINT, MATPTS, AUTDUM1 and AUTDUM2. These variables were therefore identified, *a priori*, as the most promising predictors of group membership within the range of potential discriminator variables.

A summary of the results of the multiple discriminant analysis is presented in Exhibit 9.13 below. Statistica's discriminant analysis module selected the variable FACPOINT as the sole discriminator variable, i.e. the only strong predictor of "QUA" or "REN" group membership. Wilks' *lambda* for the model is substantially closer to 1 than to 0, indicating that the discriminatory power for the model is relatively weak. The variable FACPOINT (with the lowest partial *lambda*) was clearly the strongest predictor of group membership within the variable LASTCODE (see Exhibit 9.13).

Exhibit 9.13 : Summary of multiple discriminant analysis

Discriminant Function Analysis Summary (comfu92.sta)
Step 1, N of vars in model: 1; Grouping: LASTCODE (2 grps)
Wilks' Lambda: .85393 approx. F (1,210)=35.922 p< .0000

	Wilks' Lambda	Partial Lambda	F-remove (1,210)	p-level	Toler.	1-Toler. (R-Sqr.)
FACPOINT	1.000000	.853929	35.92222	.000000	1.000000	0.00

The F and p levels pertaining to the distances between the two group centroids are shown in Exhibit 9.14 below. This evaluation of the significance of the distances between the group means indicates that there was effective separation of the "QUA" and "REN" groups under the multiple discriminant model.

Exhibit 9.14 : Evaluation of distances between group means

F-values; df = 1,210 (comfu92.sta)

	QUA	REN
QUA		35.69306
REN	35.69306	

p-levels (comfu92.sta)

	QUA	REN
QUA		.000000
REN	.000000	

Inspection of the classification matrix (Exhibit 9.15) pertaining to the model however indicates that although 80% of the cases were correctly classified by means of the classification functions, there was a major dissimilarity with regard to the classification accuracies between the two groups: approximately 98% of the "QUA" cases, but only 26% of the "REN" cases, were correctly classified. Bearing in mind that the classification of the cases upon which the model was based represents a *post hoc* classification, and thus that a poorer level of accuracy could be anticipated in applying the classification functions to new cases, the multiple discriminant model developed here appears to have little predictive potential. The total number of cases shown in the classification matrix includes all cases where there were no missing data elements for the FACPOINT variable, i.e. the only variable in the multiple discriminant model.

Exhibit 9.15 : Classification matrix under multiple discriminant model

Classification Matrix (comfu92.sta)
Rows: Observed classifications
Columns: Predicted classifications

	Percent Correct	QUA p=.74762	REN p=.25238
QUA	97.50000	156	4
REN	26.41509	39	14
Total	79.81221	195	18

The classification functions to be used in testing the classification on new cases would be derived from the tabulation presented in Exhibit 9.16 below:

Exhibit 9.16 : Classification functions under multiple discriminant model

Classification Functions; grouping: LASTCODE (comfu92.sta)

	QUA p=.74762	REN p=.25238
FACPOINT	2.5982	2.3834
Constant	-65.7839	-56.4865

Testing the classification functions on a range of values for the FACPOINT variable revealed that higher values for the "QUA" function were obtained for all cases with FACPOINT>43. Therefore, successful graduates within the cohort with faculty points who entered the BCom program with less than 44 faculty points would have been mis-classified as members of the "REN" group by means of the classification functions. Conversely, those students who entered the BCom program with more than 43 faculty points, and were subsequently excluded on academic grounds, would have been mis-classified as members of the "QUA" group by means of these classification functions.

FACULTY OF SCIENCE (BSc cohort)

Before carrying out the multiple discriminant analysis, the correlations between the independent variables in the analysis (i.e. SEXN, AUTDUM1, AUTDUM2, FACPOINT, ENGPTS and MATPTS) were examined (see Exhibit 9.17 below). These correlations were generally weak, although a moderate correlation between the variables FACPOINT (faculty points) and MATPTS (Swedish points equivalent of the matric maths symbol) was noted.

Exhibit 9.17 : Coefficients of linear correlation between independent variables

Total Correlations (sci92fu.sta)

	SEXN	AUTDUM1	AUTDUM2	FACPOINT	ENGPTS	MATPTS
SEXN	1.00	-.14	.09	.04	-.14	.22
AUTDUM1	-.14	1.00	-.30	.11	-.24	-.01
AUTDUM2	.09	-.30	1.00	-.41	.02	-.25
FACPOINT	.04	.11	-.41	1.00	.31	.60
ENGPTS	-.14	-.24	.02	.31	1.00	.13
MATPTS	.22	-.01	-.25	.60	.13	1.00

Prior to the analysis, a range of descriptive statistics pertaining to the independent variables in the model was explored. A tabulation of the variable means (and the group means for the "QUA" and "REN" levels within the dependent variable LASTCODE) is presented in Exhibit 9.18. Inspection of this tabulation shows that the group means for the independent variables FACPOINT, AUTDUM2 and AUTDUM1 appear to differ markedly; these variables were therefore considered to be likely candidates for inclusion within the multiple discriminant model. Conversely, given that predictor variables within multiple discriminant analysis are selected on the basis of maximal separation of group means, the variables SEXN and ENGPTS do not seem to be potentially strong predictor variables. Inspection of the standard deviations of the means (and the group means) indicated that there was a high level of variability amongst all of the independent variables.

Exhibit 9.18 : Group means, overall means and standard deviations of independent variables

Means (sci92fu.sta)

	SEXN	AUTDUM1	AUTDUM2	FACPOINT	ENGPTS	MATPTS	Valid N
QUA	.635135	.189189	.081081	49.70946	6.202703	6.192195	148
REN	.763158	.263158	.381579	43.13158	5.802631	5.618421	76
All Grps	.678571	.214286	.183036	47.47768	6.066964	5.997522	224

Standard Deviations (sci92fu.sta)

	SEXN	AUTDUM1	AUTDUM2	FACPOINT	ENGPTS	MATPTS	Valid N
QUA	.483027	.392989	.273886	7.931899	1.043017	1.279844	148
REN	.427970	.443273	.489002	7.003508	.783492	.992869	76
All Grps	.468071	.411245	.387562	8.228761	.979603	1.218784	224

In order to check that the underlying assumption of normality was not violated for any of the independent variables in the analysis, categorised normal probability plots were produced for

each of these variables (Exhibit 9.19a - f). Examination of these plots showed that the independent variables in the analysis were basically normally distributed; there were no major outliers which could have a large impact on the means, thereby violating the assumption of non-correlation between the means of the independent variables and their variances.

Exhibit 9.19a : Normal probability plot of the variable SEXN

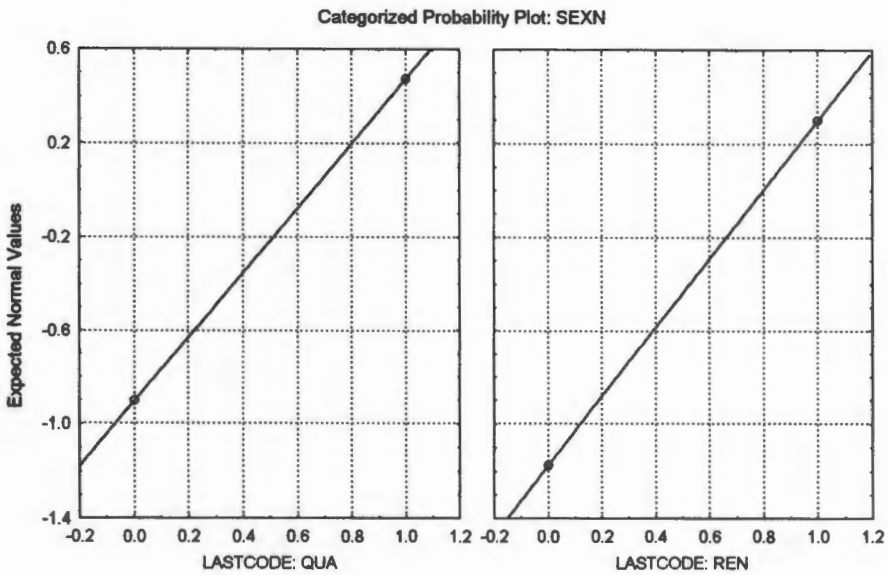


Exhibit 9.19b : Normal probability plot of the variable AUTDUM1

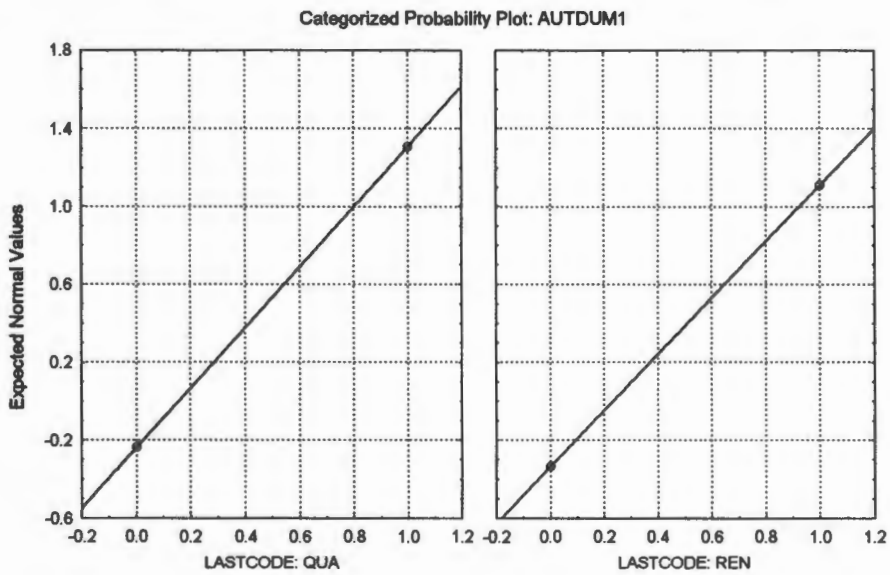


Exhibit 9.19c : Normal probability plot of the variable AUTDUM2

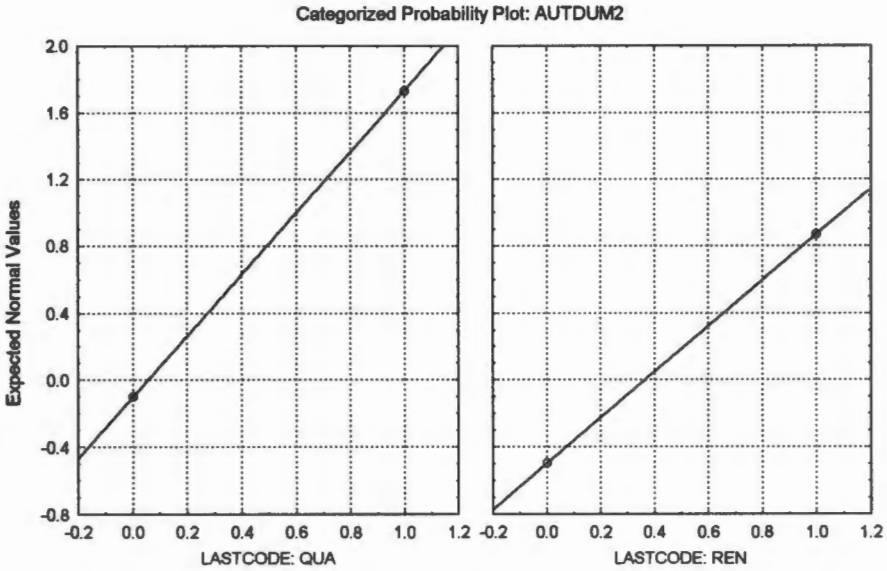


Exhibit 9.19d : Normal probability plot of the variable FACPOINT

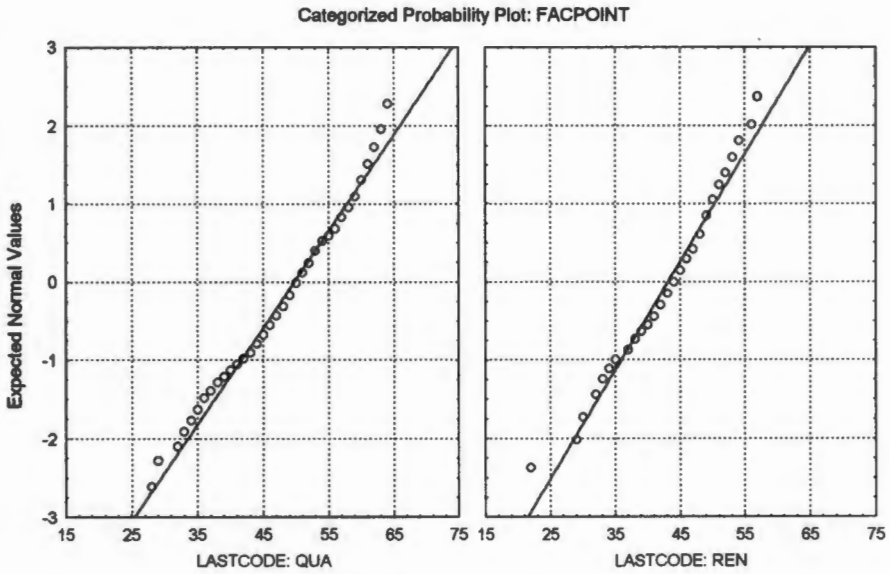


Exhibit 9.19e : Normal probability plot of the variable ENGPTS

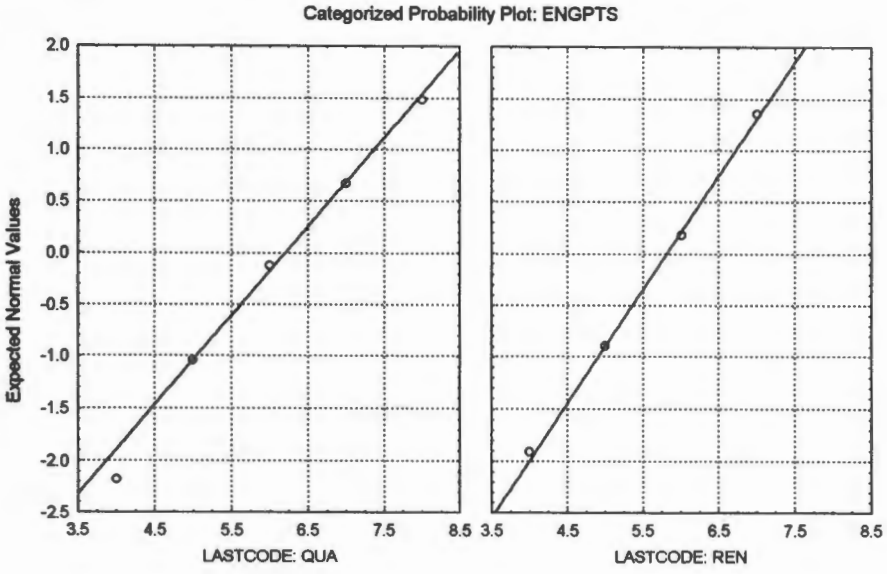
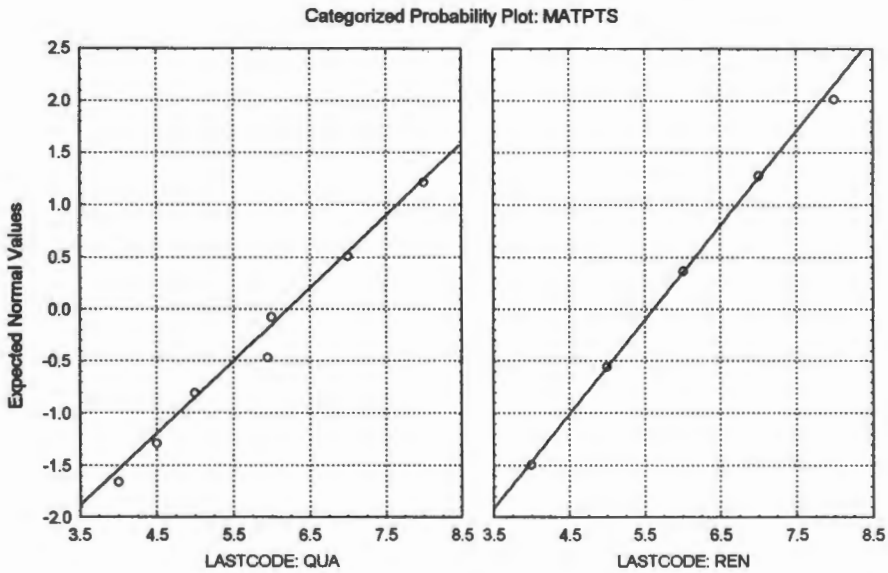


Exhibit 9.19f : Normal probability plot of the variable MATPTS



Categorised box-and-whisker plots for each of the independent variables (Exhibit 9.20a - f) served to visualise the wide dispersion of the variables about the group means, although there appeared to be considerable overlap in the 1.96*standard deviation whiskers of the two group means within the variables FACPOINT, AUTDUM1 and AUTDUM2, and possibly SEXN. This observation supports the *a priori* identification of these variables as potential predictors of group membership within the dependent variable LASTCODE.

Exhibit 6.20a : Box-and-whisker plot of the variable SEXN by group

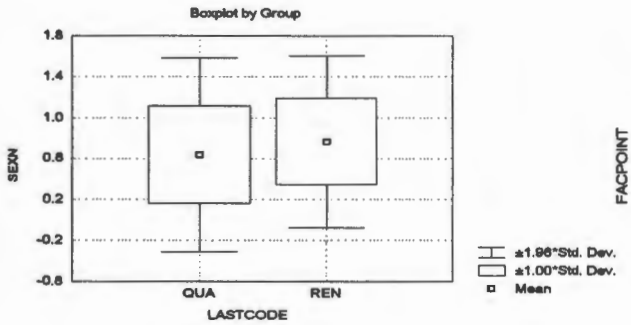


Exhibit 6.20b : Box-and-whisker plot of the variable AUTDUM1 by group

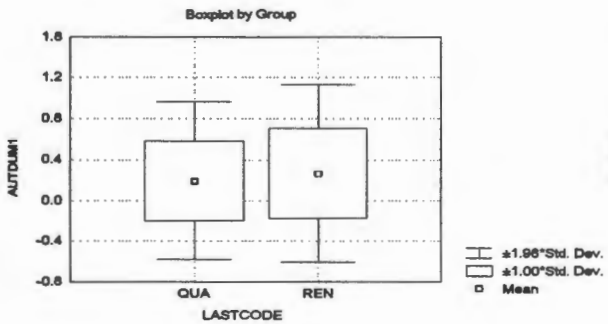


Exhibit 6.20c : Box-and-whisker plot of the variable AUTDUM2 by group

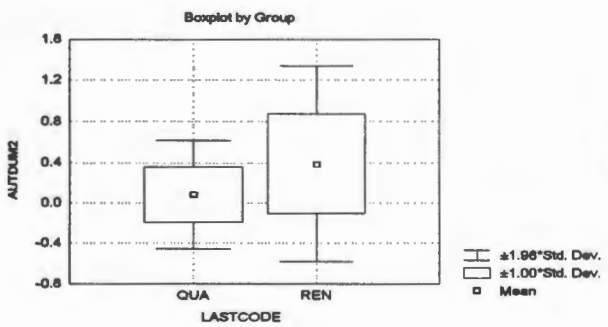


Exhibit 6.20d : Box-and-whisker plot of the variable FACPOINT by group

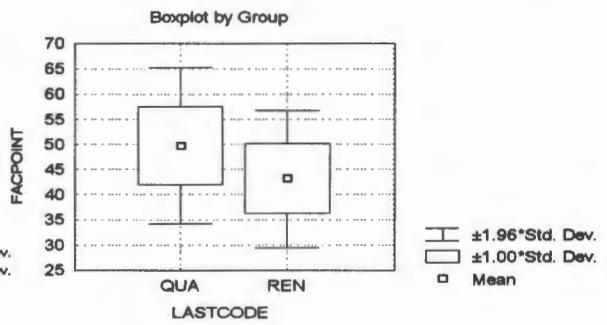


Exhibit 6.20e : Box-and-whisker plot of the variable ENGPTS by group

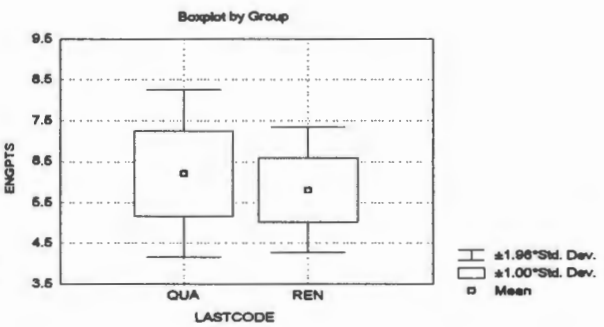
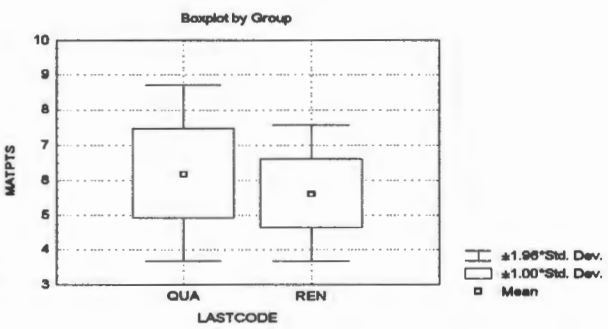


Exhibit 6.20f : Box-and-whisker plot of the variable MATPTS by group



A summary of the multiple discriminant analysis is shown in Exhibit 9.21 below. The significant p value ($p<0.000$) returned by the analysis suggests that the "QUA" and "REN" groups differ significantly with respect to the variables included in the analysis. Exhibit 9.21 indicates the that the forward stepwise process identified the variables FACPOINT, AUTDUM2, AUTDUM1 and SEXN as significant discriminators between the "QUA" and "REN" cases within a highly significant multiple discriminant model. The implication of the relatively high Wilks' *lambda* for the model (0.77064) , however, is that this model has only moderate discriminatory power. Examination of the partial lambda's for the model shows that the variable FACPOINT, which has the lowest partial *lambda,* contributes the greatest discriminatory power.

Exhibit 9.21 : Summary of multiple discriminant analysis

Discriminant Function Analysis Summary (sci92fu.sta)
Step 4, N of vars in model: 4; Grouping: LASTCODE (2 grps)
Wilks' Lambda: .77064 approx. F (4,219)=16.295 p< .0000

	Wilks' Lambda	Partial Lambda	F-remove (1,219)	p-level	Toler.	1-Toler. (R-Sqr.)
FACPOINT	.821021	.938633	14.31793	.000199	.827588	.172412
AUTDUM2	.823015	.936359	14.88465	.000150	.763919	.236081
AUTDUM1	.802949	.959759	9.18227	.002738	.895918	.104082
SEXN	.783885	.983100	3.76481	.053625	.970147	.029853

Examination of the distances between groups under the model (see Exhibit 9.22 below) nevertheless shows that there was significant discrimination between the "QUA" and "REN" groups at the 95% confidence level, when evaluated in terms of the distance between these group means.

Exhibit 9.22 : Evaluation of distances between group means

F-values; df = 4,219 (sci92fu.sta)

	QUA	REN
QUA		16.26110
REN	16.26110	

p-levels (sci92fu.sta)

	QUA	REN
QUA		.000000
REN	.000000	

Scrutiny of the classification matrix (Exhibit 9.23 below), which represents a *post hoc* classification of the cases based on which the discriminant functions were derived, shows that 76% of these cases were classified correctly using the classification functions depicted in Exhibit 9.24. The total number of cases shown in the classification matrix includes all cases where there were no missing data elements for the any of the four variables included in the multiple discriminant model. Once again, however, the classification of "QUA" cases (89% correct classifications) was seen to be far more accurate than that of "REN" cases (51% correct classifications). Given that the application of the classification functions to new cases

would be likely to result in a less accurate classification, these classifications do not appear to be of great value in the prediction of likely graduates and, more importantly, in the prediction of likely academic exclusions amongst first-time entering BSc students.

Exhibit 9.23 : Classification matrix under multiple discriminant model

Classification Matrix (sci92fu.sta)
Rows: Observed classifications
Columns: Predicted classifications

	Percent Correct	QUA p=.66071	REN p=.33929
QUA	88.74172	134	17
REN	51.28205	38	40
Total	75.98254	172	57

Examination of the classification functions (Exhibit 9.24) indicated that increasing values for the variable FACPOINT would result in higher values for the "QUA" function, and that AUTDUM2=1 (where the matric authority was DET) or AUTDUM=1 (where the matric authority was CO) or SEXN=1 (where the gender was male) would result in an increase in the value for the "REN" classification function. In summary, high faculty points appear to discriminate in favour of undergraduate success (the "QUA" group), whilst matriculation under the DET or CO matriculation authorities and, to a lesser extent, being of the male gender discriminated against ultimate undergraduate success.

Exhibit 9.24 : Classification functions under multiple discriminant model

Classification Functions; grouping: LASTCODE (sci92fu.sta)

	QUA p=.66071	REN p=.33929
FACPOINT	1.0288	.9433
AUTDUM2	9.9115	11.8256
AUTDUM1	1.8419	3.0780
SEXN	1.7546	2.4342
Constant	-27.1181	-25.0145

FACULTY OF SOCIAL SCIENCE (BA/BSocSc cohort)

In preparation for multiple discriminant analysis, a tabulation of the correlations between the independent variables in the analysis (i.e. SEXN, AUTDUM1, AUTDUM2, FACPOINT, ENGPTS and MATPTS) was examined (see Exhibit 9.25 below). These correlations were generally weak, but there was a moderate positive correlation between the variables FACPOINT and MATPTS; a moderate negative correlation between the MATPTS and the AUTDUM2 (where the DET matric authority as coded as 1) variables was also noted.

Exhibit 9.25 : Coefficients of linear correlation between independent variables
Pooled Within-Groups Correlations (sshfu92.sta)

	SEXN	AUTDUM1	AUTDUM2	FACPOINT	ENGPTS	MATPTS
SEXN	1.00	-.07	.16	-.11	-.13	-.05
AUTDUM1	-.07	1.00	-.18	.04	-.09	-.08
AUTDUM2	.16	-.18	1.00	-.47	.01	-.52
FACPOINT	-.11	.04	-.47	1.00	.41	.54
ENGPTS	-.13	-.09	.01	.41	1.00	-.03
MATPTS	-.05	-.08	-.52	.54	-.03	1.00

A range of descriptive statistics pertaining to the independent variables in the analysis was then examined. A tabulation of the means and group means, and the standard deviations of the independent variables is presented in Exhibit 9.26 below. It appears that there are marked differences in the group means within the variables SEXN, AUTDUM1 and AUTDUM2. Examination of the standard deviations of these means however indicates that there was a high level of variability within all three of these variables; this could adversely influence the separation of the groups within the dependent variable LASTCODE.

Exhibit 9.26 : Group means, overall means and standard deviations of independent variables

Means (sshfu92.sta)

	SEXN	AUTDUM1	AUTDUM2	FACPOINT	ENGPTS	MATPTS	Valid N
QUA	.340426	.164894	.132979	34.08511	6.250000	4.361702	188
REN	.565217	.304348	.086957	33.82609	5.608696	4.260870	23
All Grps	.364929	.180095	.127962	34.05687	6.180095	4.350711	211

Standard Deviations (sshfu92.sta)

	SEXN	AUTDUM1	AUTDUM2	FACPOINT	ENGPTS	MATPTS	Valid N
QUA	.475118	.372075	.340458	5.012625	.869107	1.494919	188
REN	.506870	.470472	.288104	2.774104	.891328	1.321755	23
All Grps	.482555	.385180	.334842	4.815311	.892179	1.474461	211

In order to establish that the assumption of normality was not violated within the independent variables selected for the analysis, a set of categorised normal probability plots of these variables was inspected (Exhibit 9.27a - f). It appeared that, with the possible exception of the FACPOINT variable, the independent variables were basically normally distributed and that there were no major outliers which would affect the magnitude of their means.

Exhibit 9.27a : Normal probability plot of the variable SEXN

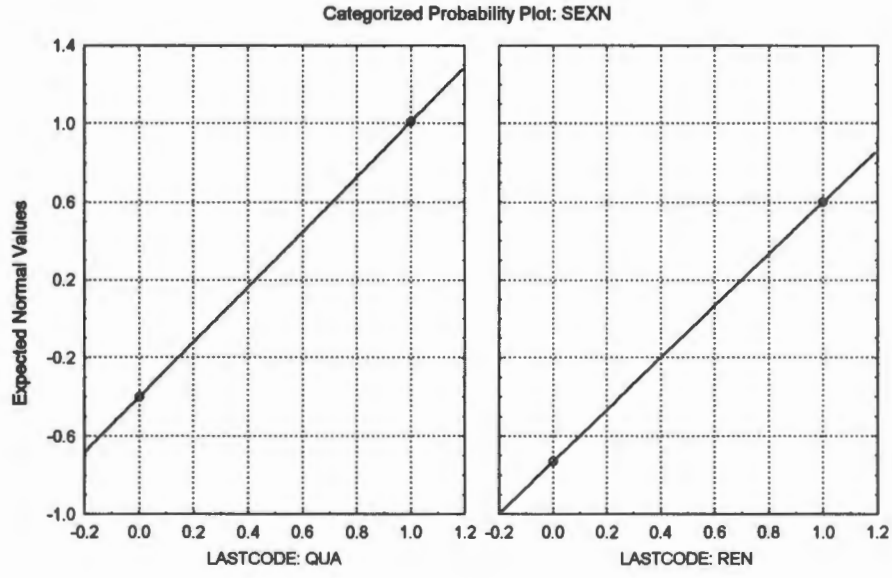


Exhibit 9.27b : Normal probability plot of the variable AUTDUM1

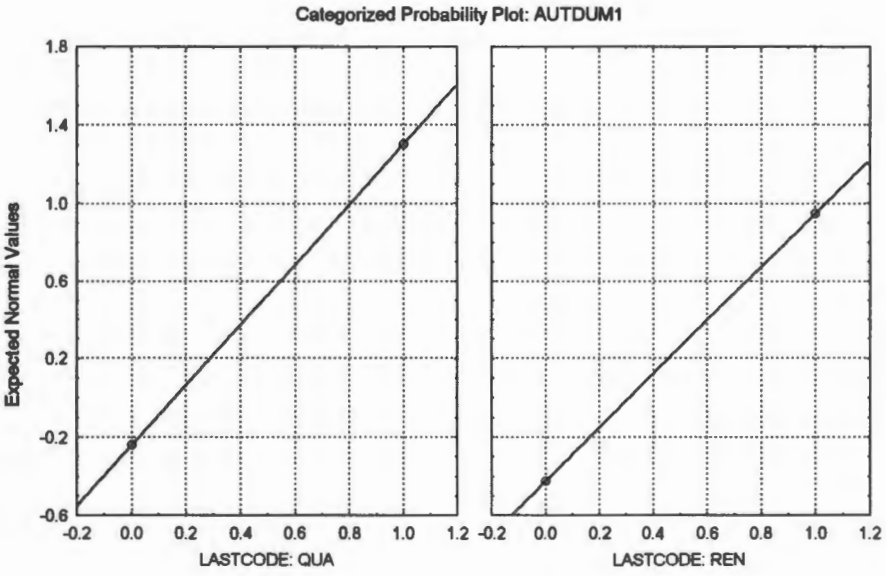


Exhibit 9.27c : Normal probability plot of the variable AUTDUM2

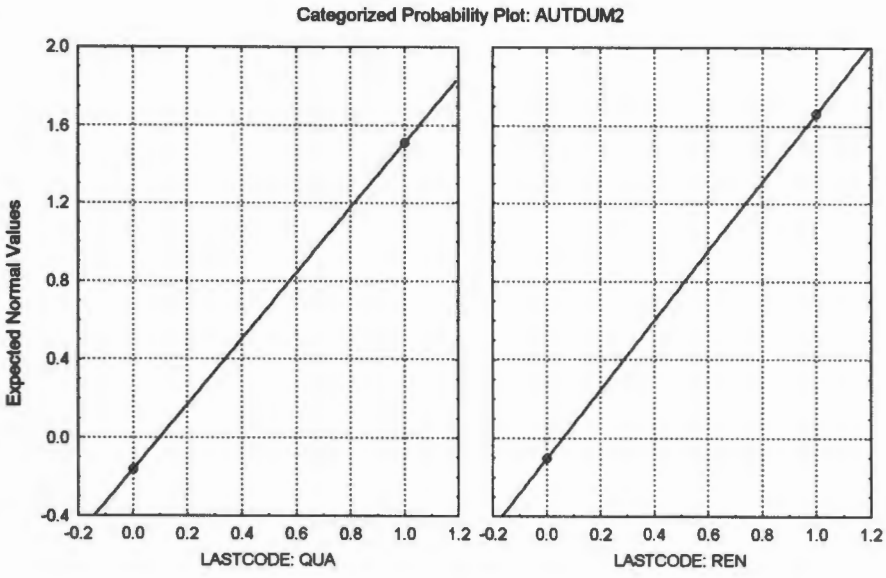


Exhibit 9.27d : Normal probability plot of the variable FACPOINT

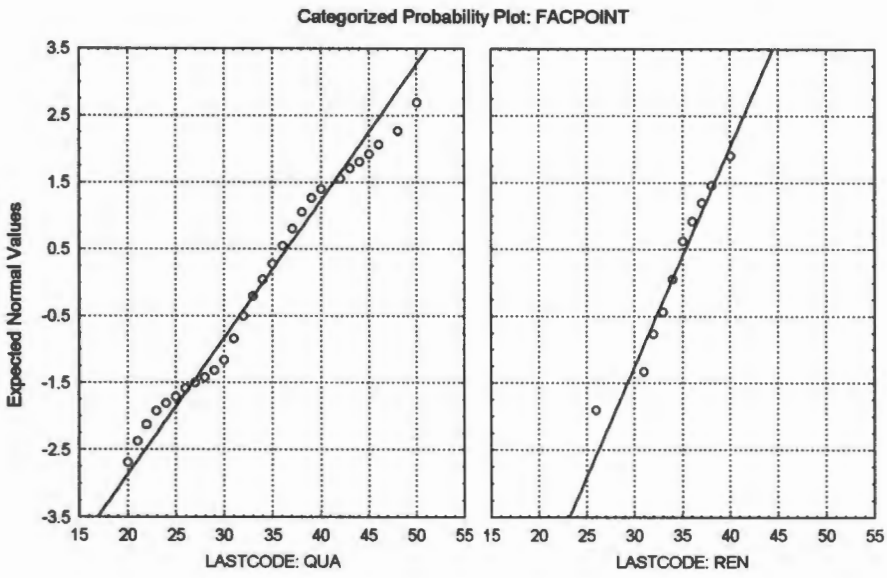


Exhibit 9.27e : Normal probability plot of the variable ENGPTS

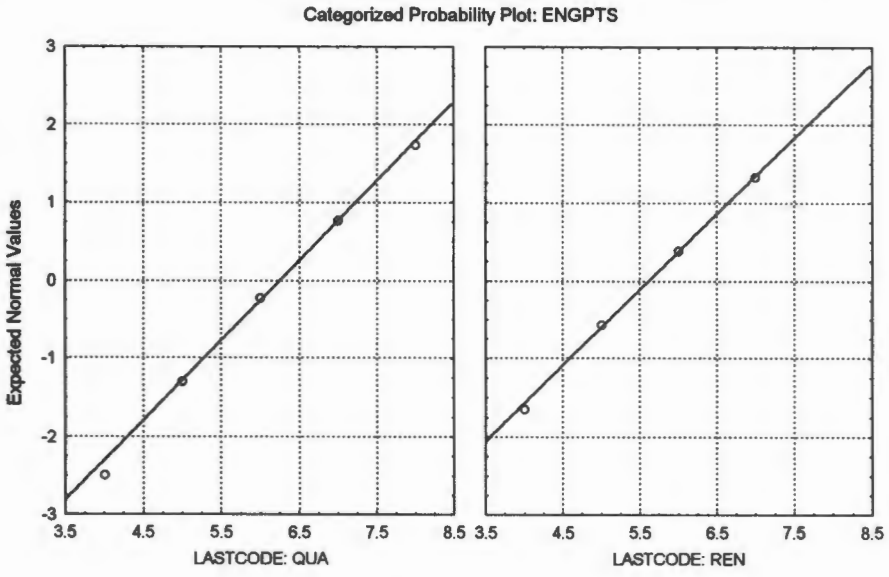
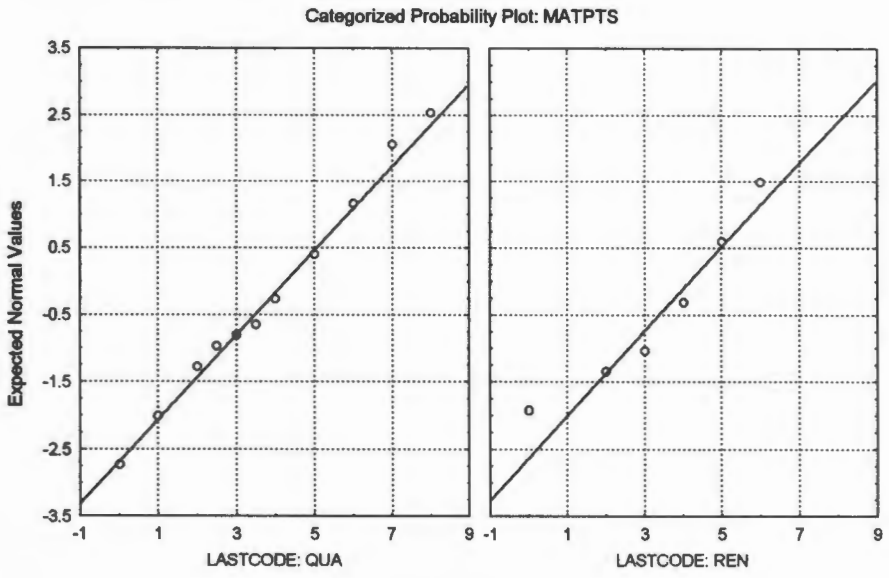


Exhibit 9.27f : Normal probability plot of the variable MATPTS



Categorised box-and whisker plots of the independent variables in the analysis (see Exhibit 9.28 a - f) confirm that the data within the independent variables selected for the analysis were widely distributed around the means; as a result, there was considerable overlap between the 1.96*standard deviations of the two group means for each variable, and in some cases the data distribution for one of the groups completely included that of the other group (ie AUTDUM1, AUTDUM2 and FACPOINT). The clearest separations between group means were apparent within the variables SEXN (the binary coded gender variable) and ENGPTS (representing the Swedish points equivalent of the matric English symbols achieved).

Exhibit 6.27a : Box-and-whisker plot of the variable SEXN by group

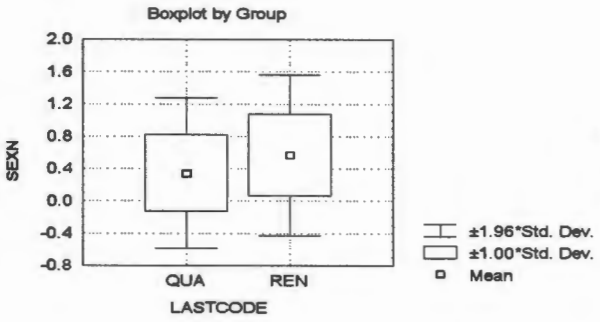


Exhibit 6.27b : Box-and-whisker plot of the variable AUTDUM1 by group

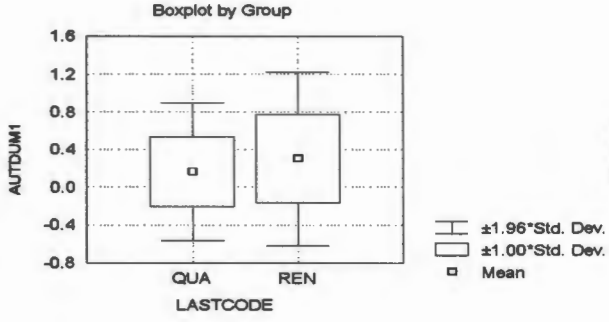


Exhibit 6.27c : Box-and-whisker plot of the variable AUTDUM2 by group

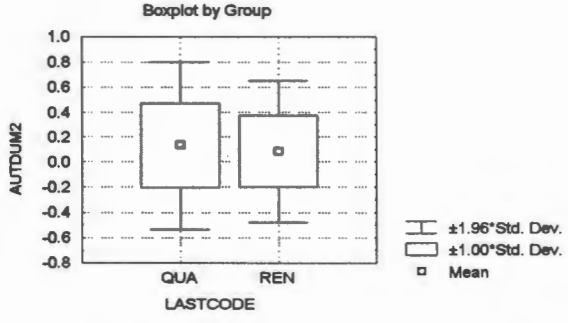


Exhibit 6.27d : Box-and-whisker plot of the variable FACPOINT by group

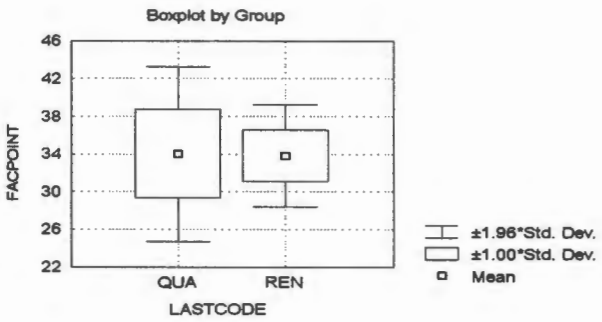


Exhibit 6.27e : Box-and-whisker plot of the variable ENGPTS by group

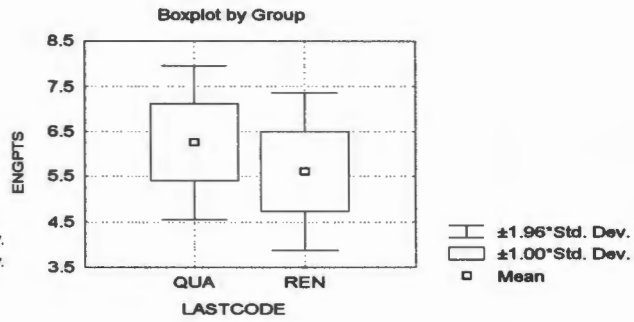
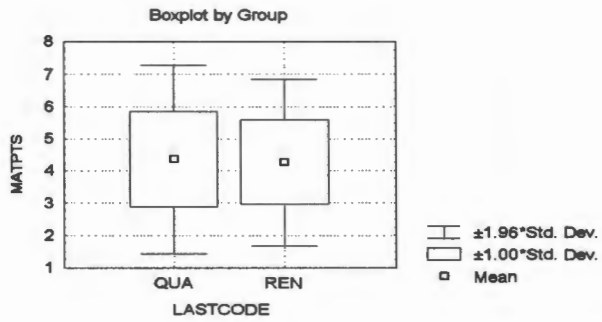


Exhibit 6.27f : Box-and-whisker plot of the variable MATPTS by group



A summary of the results of the forward stepwise discriminant analysis (see Exhibit 9.29) indicates that there was a significant difference between the two groups within the dependent variable LASTCODE ($p<.05$). Wilks' *lambda* for the model (1.000) however established that the model had no discriminatory power, and it appeared that the classification matrix and classification functions pertaining to this model would therefore be of little value. In summary, multiple discriminant analysis selected the variable ENGPTS as the sole significant predictor of "QUA" or "REN" group membership within the variable LASTCODE amongst the matric performance-related independent variables in the analysis.

Exhibit 9.29 : Summary of multiple discriminant analysis

Discriminant Function Analysis Summary (sshfu92.sta)
Step 1, N of vars in model: 1; Grouping: LASTCODE (2 grps)
Wilks' Lambda: .94958 approx. F (1,209)=11.097 p< .0010

	Wilks' Lambda	Partial Lambda	F-remove (1,209)	p-level	Toler.	1-Toler. (R-Sqr.)
ENGPTS	1.000000	.949579	11.09745	.001022	1.000000	0.00

Examination of the distances between groups under the model (see Exhibit 9.30 below) nevertheless shows that there was significant discrimination ($p<.05$) between the "QUA" and "REN" groups at the 95% confidence level, when evaluated in terms of the distance between these group means.

Exhibit 9.30 : Evaluation of distances between groups under the multiple discriminant model

F-values; df = 1,209 (sshfu92.sta)

	QUA	REN
QUA		10.76154
REN	10.76154	

p-levels (sshfu92.sta)

	QUA	REN
QUA		.001214
REN	.001214	

Inspection of the classification matrix (Exhibit 9.31) revealed that whilst 100% of the members of the "QUA" group were correctly classified by means of the *post hoc* classification of cases, all of the "REN" cases were mis-classified as "QUA" cases. The classification functions (Exhibit 9.31) are therefore of little predictive value in forecasting the probable "QUA" or "REN" group membership of new cases. The total number of cases shown in the classification matrix includes all cases where there were no missing data elements for the ENGPTS variable, which is the only variable in the multiple discriminant model.

Exhibit 9.31 : Classification matrix under multiple discriminant model

Classification Matrix (sshfu92.sta)
Rows: Observed classifications
Columns: Predicted classifications

	Percent Correct	QUA p=.89100	REN p=.10900
QUA	100.0000	248	0
REN	0.0000	33	0
Total	88.2562	281	0

Testing of the classification functions for "QUA" or "REN" group membership (Exhibit 9.32) revealed that the value for the "QUA" function was higher than that for the "REN" function for all values for the ENGPTS variable greater than 3. All of the "REN" cases within the cohort had ENGPTS values of 4 or more, and hence these cases were all mis-classified as members of the "QUA" group under the discriminant functions above.

Exhibit 9.32 : Classification functions under the multiple discriminant model

Classification Functions; grouping: LASTCODE (sshfu92.sta)

	QUA p=.89100	REN p=.10900
ENGPTS	8.2295	7.3851
Constant	-25.8325	-22.9266