1980

Forecasting enrollments in a Virginia community college

Sue C. Lawrence
College of William & Mary - School of Education

Recommended Citation

Lawrence, Sue C., "Forecasting enrollments in a Virginia community college" (1980). *Dissertations, Theses, and Masters Projects.* Paper 1539618573.
https://dx.doi.org/doi:10.25774/w4-gd66-bc63

INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.
2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.
3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.
4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.
5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University
Microfilms
International

300 N. ZEEB ROAD, ANN ARBOR, MI 48106
18 BEDFORD ROW, LONDON WC1R 4EJ, ENGLAND

8103584

LAWRENCE, SUE CHAMPNEY

FORECASTING ENROLLMENTS IN A VIRGINIA COMMUNITY COLLEGE

The College of William and Mary in Virginia ED.D. 1980

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

FORECASTING ENROLLMENTS
IN A VIRGINIA COMMUNITY COLLEGE

A Dissertation Presented to The Faculty
of The School of Education of
The College of William and Mary in Virginia

In Partial Fulfillment of the Requirements
for the Degree Doctor of Education

---

by Sue C. Lawrence
July 1980

FORECASTING ENROLLMENTS
IN A VIRGINIA COMMUNITY COLLEGE

by

Sue C. Lawrence

Approved July 29, 1980 by

Robert B. Bloom, Ph.D.

Clifton Conrad, Ph.D.

D. J. Herrmann, Ph.D.,
Chairman of Doctoral Committee

ACKNOWLEDGEMENTS

I would like to express my appreciation to Dr. O. J. Herrmann, who directed this research, for his advice and encouragement, and to Drs. Clifton Conrad and Robert Bloom who served as members of my dissertation committee, for their assistance and guidance.

I would also like to express my appreciation to Thomas Nelson Community College (TNCC) for granting me a leave of absence during the 1978-79 academic year. Dr. Stuart Bounds, Director of Institutional Research, and Mrs. Nancy Merriner, Statistician at TNCC, rendered invaluable aid in collecting data and gathering information.

Lastly, I would like to express my appreciation to my husband, James D. Lawrence, Jr., for his aid and encouragement, and to my mother, Mrs. Inez Champney, for her encouragement. Without their continued support this project would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: THE PROBLEM

Introduction

Since the return of the veterans following World War II, institutions of higher education in the United States have been in a period of accelerated growth, both in terms of enrollment and public support. It now appears that growth may be ending and that enrollment patterns may change. There are also indications that the public generally and members of legislatures particularly are becoming more conservative in the support provided for institutions of higher education (Cheit, 1971). One result of these changes has been the demand for accurate projections of student enrollment in colleges and universities, primarily for the purposes of fiscal planning. The problem of accurately projecting enrollments has become particularly acute in institutions with open admission policies such as community colleges. Many attempts have been made to develop accurate forecasting models. Models such as the cohort survival, the student flow using Markov transition matrices, multiple regression, nonlinear growth, time-series decomposition, surveys, and others have been used to predict enrollment. It appears that different techniques have worked in different circumstances, and that each individual institution must develop its own model, applying its own special characteristics.

Community college enrollments pose difficult enrollment forecasting problems because of characteristics unique to these institutions. The community college is a multidimensional, open admissions institution oriented toward meeting the needs of a particular community. They serve all citizens in one geographical area called the service area. This area is usually designated by the State. The student body is not composed of a well defined group of students, but includes students of many ages, races, sex, and levels of academic preparation. In addition to the traditional college population, adults, minorities, slow learners, students with academic deficiencies, and the elderly are attending community colleges (Harrington, 1977). Because of the diversity of the student population and the different priorities of these students, it has been difficult to find stable trends in enrollment.

Changing enrollment patterns have increased the difficulty of forecasting enrollments. Probably the largest single problem in projecting enrollments for community colleges results from the increase in part-time student enrollment (NCES, 1978). Since part-time student enrollment is increasing much more rapidly than full-time student enrollment, headcount enrollment is increasing much faster than full-time equivalent enrollment (FTE). FTE enrollment is the basis on which states, including Virginia, allocate funds for their state colleges and universities. The large increases in part-time student enrollment are due primarily to the increase in the number of older students. In 1975, for example, of the 18 and 19 year old students attending college, 92 percent were full-time; the percentage of older students, on the other hand, attending full time was 23 percent (NCES, 1978). NCES (1978) predicts this pattern will change because the population projections for the years 1977 through 1986 indicate that the number of 16 to 22 year olds will decrease, and the number of older students will increase.

A second change in enrollment patterns which has affected the community colleges, and their full-time equivalent enrollment is the proportion of female students. In community colleges especially, female enrollment has increased very rapidly, and by the end of 1980 may become greater than male enrollment (NCES, 1978). Finally, increasing numbers of minorities and disadvantaged students are attending community colleges. For these new nontraditional students in higher education, new programs are required. In order to meet the needs of these nontraditional students, classes are scheduled in the evening as well as during the day, and are taught off campus as well as on campus. These students have required refresher courses and new technological advances in teaching methods. As a result the cost per student has risen (Harrington, 1977). The increase in the cost per student, coupled with the decrease (percentage wise) in total public support, requires community colleges to project their enrollments accurately. In many states, such as Virginia, colleges are financially penalized for over or under projecting their enrollments (Virginia Appropriation Act, 1978).

Need for the Study

During the 1960's the community college movement experienced its greatest period of growth, so there was little need or incentive to develop accurate procedures for forecasting enrollment. In the late 1960's, however, several changes took place causing community colleges and other institutions of higher education to investigate forecasting procedures. As inflation increased, costs began rising while enrollments tended to decline (Cheit, 1971). As the general public became increasingly concerned with high taxes and inflation, they began questioning the value and, in turn, the cost of higher education. State legislatures came to examine college budgets and to demand more specific many services had to be reduced or eliminated until the next budget cycle. Since the legislature penalizes the institution for both over projections and under projections, it is clearly designed to limit "gamesmanship" by the institutions in their quest for State funds. It is also evident that this legislation punishes the institutions for failing to submit accurate enrollment projections by requiring them to return funds already committed for faculty salaries or other operating expenses if the projections are not accurate. Thus, for the state supported institutions of higher education who pursue an open door admissions policy, accurate forecasting of student enrollment presents an especially difficult problem.

As the need for more accurate enrollment predictions became critical, it also became obvious that the methods that had been used in making these projections were inadequate. Through 1970, higher education was a growth industry (Bowen, 1974). Projections for enrollment were made by simply assuming the continuance of the growth trend. In 1970, the growth rate of enrollments began to decrease and, in 1976, enrollments actually fell (see figure 1, Appendix I). New techniques were needed to provide accurate projections of enrollment, both short and long term, in order to optimize the expenditures of available funds and the utilization of limited resources.

Purpose

The purpose of this study is to identify the factors relevant for projecting enrollments and to develop such factors into a model that will project enrollments within ±1 percent accuracy for a public community college using expo facto data from a Virginia community college and justification for projected expenditures. In Virginia, for example, budgetary requests for state colleges and universities must now be justified on the basis of accurately projected enrollments. In its Appropriation Act of 1978, the General Assembly of Virginia gave the State Council of Higher Education the responsibility for developing "policies, formulas, and guidelines for the fair and equitable distribution and use of public funds among the institutions of higher education, taking into account enrollment projections" (Code of Virginia, 1978, p. 4).

As a result of this legislation, an institution may be required to return appropriated funds in excess of those justified by the actual enrollment to the Commonwealth. Thus, if the institution's enrollment exceeds the projected enrollment by 1 percent or more, "the Governor is authorized to direct the transfer to the surplus of the general fund, from the appropriation to that institution for educational and general activities, an amount not exceeding the tuition and fees collected on account of the enrollment in excess of 1 percent. However, the transfer shall not be made for any excess in enrollment which is less than 50 full-time equivalent students" (Virginia Appropriation Act, 1978, p. 166). According to policy, leniency should be allowed for colleges having a large proportion of part-time students (Virginia Appropriation Act, 1978). In 1976, however, Thomas Nelson Community College was required to return $44,000 to the Director of Budget as a result of the over projection of enrollment in spite of a heavy proportion of part-time students in the student population. Most of these funds had either been committed or already expended for faculty, personnel, programs, or services before the order to return some was received. As a result, utilizing the data format developed by the State Council of Higher Education in Virginia.

Hypotheses

If the purposes of this study are to be achieved, it is necessary to test the following hypotheses:

1. Variables can be identified that will produce accurate forecasting models within the limits of accuracy designated by the Virginia State Legislature.
2. Using the variables identified above, forecasting models can be developed for projecting the following categories of enrollment:
    a. Fall headcount and summer session headcount within ±1 percent, and
    b. Total FTE within ±1 percent.

Definition of Terms and Abbreviations

Since this study involves two fields, higher education and statistical forecasting, it may be helpful to define several terms which may not be familiar to researchers in both fields. Definitions and abbreviations of various terms used in this study are given in the sections to follow:

1. Headcount Enrollment includes every student taking a course without regard to the number of credit hours taken.
2. Full-time Equivalent Enrollment (FTE) equals the total credit hours taken collectively by the students divided by 15 (the number of credits defined as a full undergraduate student lot in the quarter system).

3. Open Admissions Institutions are those institutions who accept all qualified students who desire to attend and for whom programs are available.
4. College Going Rate is the ratio of the number of students going to a particular college divided by the total population in the age group from which the students come. A synonym is Penetration Ratio.
5. Persistence Rate or Survival Rate is the percent of an original group that returns for the next stage after completing a previous stage of activity.
6. Participation Rate is the percentage of a particular group, such as a given population, that participates in a given activity.
7. A Cohort is a group of individuals having a common classification or trait.
8. Universal Access to Higher Education means "the guarantee of a place for every high school student who wishes to enter higher education" (Carnegie, 1973).
9. Relative Educated Wage is the wage of educated relative to uneducated labor (Dresch, 1975).
10. The Active Adult Population (AAP) is the population between the ages of 25 and 64 (Dresch, 1975).
11. The Transverse Progression Theory says that a particular population grows exponentially.
12. Endogenous Data are internal data generated because the system exists.

12. Endogenous Data are internal data generated because the system exists.
13. Exogenous Data are data external to the system; the system has no effect on them.
14. Regressor Variables are the independent variables in a regression equation.
15. A Global Model is one in which the structure is highly stable and the chosen model is the truth about the underlying structure of the data (Gilchrist, 1978).
16. A Local Model is one in which the structure is stable for the short run, but not necessarily in the long run (Gilchrist, 1978).
17. Deterministic Models are perfect mathematical forms with no chance element; i.e., the forecast of the next observation would be a perfect forecast.
18. Stochastic Models are mathematical forms where the chance element plays a major role.
19. Explicit Form is the mathematical form when all data are used in the forecast.
20. Recurrence Form is a mathematical form in which the new forecast can be obtained from the old forecast with a minimum amount of effort when new observations are added.
21. Forecast Error Form is a mathematical form in which a multiple of the last forecast error is added to the last forecast to obtain the new forecast.
22. Leading or Lagged Variables refers to terms shifted forward or backward in time by one or more years or quarters.

24. LEI is a composite of 12 leading economic indicators computed by NBER and published in the Business Conditions Digest.
25. NBER - National Bureau of Economic Research.
26. NCES - National Center for Education Statistics.
27. $\bar{e}$ - the value of x for which $\ln x = 1$.
28. TNCC - Thomas Nelson Community College.
29. VCCS - Virginia Community College System
30. SCHEV - State Council of Higher Education of Virginia.
31. SBO - State Budget Office.
32. SREB - Southern Regional Education Board

Plan of Study

In Chapter 1 the problem to be studied--forecasting enrollments at a community college--was introduced along with the need for the study, the purpose, and hypotheses. Chapter 2 reviews the relevant literature, explores the different forecasting techniques, and gives examples where each technique has been applied. The procedures used to develop an enrollment forecasting model for a community college is given in chapter 3. A Virginia community college was used as an example. Although the general procedures for developing a forecasting model are the same, each model must be particularized to an individual college because of the difference in characteristics. The optimum forecasts and the models producing the optimum forecasts are discussed in chapter 4 along with the relevant statistics. Summaries, interpretations, conclusions, and implications for further studies are discussed in chapter 5.

CHAPTER 2: PREVIOUS RESEARCH

Introduction

In the last decade the problem of projecting enrollments at all levels has been studied and reported in detail. Most institutions of higher education are interested in enrollment projections because they are closely related to institutional goals and purposes and are essential to financial and program planning at every level. If the resources available to institutions of higher education fluctuate in the future as predicted, the problem will become of increasing importance.

This chapter is organized into five major topics: the scope of the problem, changing patterns in enrollment, the factors relevant to this study, the accuracy of past enrollment projections, and a review of the methodology currently being used to project enrollments. The first section reviews current and previous enrollment projections at the national level. The second section reviews changing patterns of enrollment. In the third section, three general categories of factors relevant to projecting enrollments are discussed: demographic, economic, and institutional. After the section on relevant factors, a brief discussion is also included on the accuracy of national enrollment projections. The concluding section on methodology includes a review of current methods of statistical forecasting and research design. State and institutional studies are used to illustrate each research design.

The Scope of the Problem

In the 1970's the demands on higher education have increased. The general call for "accountability" has caused educators to place much more emphasis on systematic planning. Because enrollments determine income from tuition and fees as well as funding from public sources, enrollment projections have become an essential element in financial and program planning for institutions of higher education. In years prior to 1970, the projection of total enrollments at the national and state level was all that was required because growth rates justified the need for increased resources. In the seventies, however, growth patterns for institutions of higher education became uncertain and new trends were not easily identifiable. The large body of literature written in the seventies suggest a range of opinion on future enrollments in higher education.

At the national level, a number of enrollment projections have been made. Bowen (1974) foresees a high probability of enrollments in higher education doubling or tripling before the end of the decade. The Carnegie Commission predicts small decreases in enrollments in the early 1980's and a subsequent period of stability in enrollments to the year 2000. Dresch (1975) and Froomkin (1974) predict fairly large decreases in enrollments in the early 1980's. Cartter, "known for his record of 'most accurate' predictions", sees a decline in enrollments in the early 1980's and a modest recovery in the late 1980's (Leslie and Miller, 1974, p. 10). The National Center for Education Statistics (1978) diffuses its predictions by making three projections: high, low, and intermediate alternatives. In the NCES low alternative projection, slight increases are seen until 1981 and then steady decreases until 1985. In its high alternative, steady increases are predicted to the year 1985 (Centra, 1980).

The large variations among enrollment projections occurred because the forecasts are based on different assumptions. The assumptions differ primarily because the forecasters have diverse opinions on the future of higher education. In the past, enrollment projections were approached from the viewpoint that the past was indicative of the future. A trend-demographic approach to enrollment projecting was widely used because enrollments were assumed to grow as the population of the 18 to 22 year olds grew. In the early 1970's the growth rate of the 18 to 22 year olds began to decline (Carnegie, 1973). At the same time that falling enrollments were projected, signs of financial stress began to be evident (Cheit, 1971). Institutions of higher education began reexamining their purposes and goals which were reflected in their enrollment projections. In the following paragraphs, several major enrollment projections are reviewed in light of their underlying assumptions.

Howard Bowen, the most optimistic of national forecasters, forecasts a 200-percent increase in enrollments by the year 2000 (Carnegie, 1975). He bases his prediction on the assumption that the growth of higher education enrollments will parallel the growth in the service sector of the economy (it increased from 28 percent to 60 percent from 1900 to 1960). Bowen contends that there are a host of factors other than demographics which could affect the number of students attending institutions of higher education. These might include the number and kind of institutions available, the relevancy and attractiveness of the program, the convenience of the times and places at which education is offered, the nature of the admission requirements, the tuition charges, the terms of financial aid, and arrangement for release time from work for education. He contends that demographers who project enrollments are not inept, but human values and behavior are subject to seemingly unexplainable shifts. Bowen believes that if the political and educational leadership in this country will display more imagination and daring, enrollments will continue to rise as they have in the past (Bowen, 1974). The expansion possibilities of higher education are certainly considerable. In 1973-74, for example, only 43 percent of the total enrollment were female; in 1972 only 15 percent of the persons enrolled from the 18 to 24 year old age group were from very low income families (with less than $3000 total income). Moreover, in some states there are wide variations in college attendance. Georgia, for instance, had only 24 percent of its population in the college age group attending college in 1970 compared to Arizona with 59 percent.

The Carnegie Commission based their original projection on five basic assumptions: increasing high school graduation rates; increasing enrollment rates of 18 to 19 year olds, especially white men; increasing degree-credit enrollment (nondegree-credit enrollment was not considered in this projection); increasing graduate resident enrollment; and increasing first-professional degree enrollment. Since several of the assumptions proved to be incorrect, the Carnegie Commission revised their original projections in 1973. The following assumptions were changed: high school graduation rates were adjusted downward slightly, postbaccalaureate training was assumed constant at the 1969 level, Series E Census Bureau data was used instead of Series D for 1990 and 2000, and nondegree-credit enrollment was taken into account (Carnegie, 1973). In 1975 the Carnegie Commission revised its national enrollment projections for the third time. The new baseline projections assume rising enrollment rates based on past trends for the following categories: part-time students, nondegree-credit students, students 22 years of age and older, graduate and first professional degree students, women students, and minority students. Finally, it also bases its projections on the following judgments: white male enrollment rates returning to their peak levels by the year 2000, the Series F Census Bureau population projections may prove to be more reliable than Series E, the demand for schoolteachers will reflect current pupil-to-teacher ratios, the draft will not be reinstated, and student aid will continue to increase (Carnegie, 1975).

The NCES gives three groups of projections which they term a low alternative projection, an intermediate alternative projection, and a high alternative projection. Three changes in enrollment patterns have caused them to diffuse their forecasts: the distinction between degree-credit student and the nondegree-credit student is so cloudy that they have combined the two categories; the percentage of full-time and part-time students has been changing; and, finally, the average age of the college population is changing. Because the average age of college populations is changing, enrollment data were obtained from the U.S. Bureau of Census for age groups 16 to 24, 25 to 29, and 30 to 34, and by sex from 1967 through 1976. In the low alternative projection, age-specific enrollment rate was assumed to remain constant at the average of its 1975 and 1976 rate; a modest increase was assumed for part-time students. The high alternative projection was based on the age-specific enrollment rate from 1967 to 1976 continuing through 1986, and assumed a large increase in part-time students. The intermediate alternative used the average of the age-specific enrollment rates for the high and low alternatives and also assumed an average increase in part-time students. Between 1976 and 1986, NCES predicts a 21-percent increase in its low alternative enrollment forecast, a 54-percent increase in its intermediate forecast, and a 90-percent increase in its high forecast. They make no forecast beyond 1986.

Cartter (1973) admits that he has trusted the past too much in his predictions. He based his first estimates on the original U.S. Bureau of Census population projections. Using the latest Census population projections (Series F), he predicted a small increase of one and a third million students by 1980 and a drop of 2.9 million by 1988. Other assumptions upon which he bases his predictions are: 1) That there will be a continuing increase in the proportion of the age group that graduates from high school (this has dropped about 4 percent below what he expected); 2) that the proportion of 18 to 24 year olds going to college would increase by 2.5 percent (it decreased by 4 percent); and, finally, that the actual number in college would increase by 10 percent by 1980. He does, however, qualify his predictions by stating that his estimates could be quite incorrect if the learning society should develop.

In most previous national forecasts, the major assumptions are based on demographic factors. As Folger (1974) suggests, most enrollment forecasts are based on a simple ratio of enrollment to college-age population or to high school graduates. He contends that more sopisticated models should be considered because the planning needs of higher education are complex. Increased student aid and a reduced job market, coupled with the political pressures to give equal opportunity to women and minorities, have altered the mix of traditional students attracted by institutions of higher education. Folger believes an inadequate data base has not produced sufficient information on which to base accurate projections. Data, for example, on postsecondary vocational, trade, and technical education systems have had low priority compared to data on traditional college programs. The lack of data in these areas indicates to Folger that higher education has not yet committed itself to planning and management which must relate directly to the very goals of higher education. Leslie and Miller (1974) concur that the current data bases are inadequate for precise predictions, but they believe that knowing what students do now may not have any bearing on the future. Leslie and Miller, Froomkin, and Dresch take a different approach to projecting enrollments by assuming that economic conditions are related to enrollments and the past is not indicative of the future. In the following few paragraphs samples of such enrollment projections will be reviewed. Following these projections, changing enrollment patterns will be discussed.

Leslie and Miller (1974) see enrollments climbing exponentially. They conclude that the period of fluctuation which has existed in the 1970's, which they term "no growth" or the "steady state", will only be momentary. After observing the total growth picture from 1890 to 1970, they have concluded that enrollments have grown exponentially upward, which is called by economists "transverse progression". Economists theorize that certain systems, such as the growth of the economic system and the Gross National Product (GNP), follow a line of transverse progression, sometimes oscillating above or below this line. In the classical economic model, the economy is seen as a series of such oscillations which are continually correcting themselves to the line of transverse progression. Leslie and Miller (1974) hypothesize that higher education is analogous to our economic system, and the transverse progression curve, therefore, is descriptive of the higher education system. In fact, the higher education system grows as the economic system grows. When such a system falls into a "no growth" period, the system must revitalize itself. To accomplish this, certain functions must be performed by the system in order to resume its natural state of growth along the line of transverse progression. Possible functions are: the introduction of a new good or a new grade of good already in use; the introduction of a new method of production such as new labor saving machinery; the opening of new markets; the employment of a new source of supply production factors; and, finally, the reorganization of an industry, several industries, or part of an industry; i.e., a monopolization of some industry. Leslie and Miller (1974) contend that these five functions are being applied to higher education by the addition of new degrees such as the associate degree in community colleges and the external degree, by the use of media in higher educacion, by the addition of new clientele, by using new financial sources such as the federal government, and, finally, by the use of consortia to form research institutes and reduce duplication. As long as society and the economy continue to grow, Leslie and Miller predict that higher education will continue to grow. Using the transverse progression theory as their basis, they predict a 150-percent increase in enrollments by the year 2000.

Froomkin (1974) constructs three scenarios which might be classified as intuitive approaches. He bases his forecast on the availability of jobs for which college credentials are believed necessary. When the number of college graduates exceeds the number of jobs available for college graduates, enrollments will decline. Conversely, enrollments will increase when the availability of jobs increases. In Scenario 1, "Continuation of the Status Quo", Froomkin predicts that the undergraduate enrollment (including nondegree students) trends of the 1970's will continue until 1985 with the number of graduates increasing between one quarter and one third. He believes, however, that the levels of enrollment resulting from his first scenario are much too high in relation to the needs of the economy for college graduates. In Scenario 2, "Small Decline", he predicts a 20 percent decline in enrollment from those predicted in scenario 1. He bases this decline on the percentage of high school graduates enrolling in college. Even this decline will not balance the supply of jobs for college graduates with the demand for jobs. Scenario 3, "Floor Set by Labor Market Demand", is the most pessimistic of Froomkin's scenarios. Here, he predicts a 50 percent decline in enrollments from the 1974 level. This is the decline needed if supply and demand are balanced; i.e., the number of jobs for college graduates equals the number of graduates seeking jobs (Carnegie, 1975).

In a demographic-economic model using both cohorts and ratios, Dresch (1975) predicts the greatest decreases in national enrollments assuming a saturated market condition. He forecasts enrollments only for 24 year olds--not total enrollments. Between 1970 and 1980 his model predicts a 40 percent increase in enrollments which he claims is with NCES and the Carnegie Commission. From 1980 to 1990 Dresch predicts a 40-percent contraction, and between 1990 and 2000 the model predicts a further 12-percent contraction. The basic assumption underlying this model is that the educational characteristics of the active adult population (AAP), which is defined as the population between the ages of 25 and 64, can only be altered in two ways: through educational attainments of entering (25 year old) cohorts and through the educational characteristics of persons leaving the AAP (through death or retirement). Educated members of the AAP are defined as those persons completing 4 or more years of college. The educational attainments of the AAP are fixed at age 25 and no change in educational characteristics are permitted after this age. A second assumption is that the relative educated wage is inversely related to the ratio of educated to uneducated members of the AAP at any time. The relative wage is assumed to adjust instantaneously to any change in the educational composition of the AAP. Finally, the educational characteristics of entering cohorts are determined by educational wage differentials which they observed in prior periods, specifically between the ages of 17 and 24, and by the rate at which preceding cohorts were educated. Dresch realizes his model is limited since it only applies to 24 year olds. What he is attempting to show is that educational forecasting has relied too heavily on past growth patterns and has not considered other factors such as economic factors. Because of World War II and the postwar baby boom, educational institutions have grown at a much faster rate than they were growing before this period. In a sense the system is simply stabilizing itself. It has overexpanded because of an abnormal growth rate.

Continued growth has led to institutional structures in education that are highly dependent on growth for their existence and their dynamic effectiveness. Many forecasts such as those made by the Carnegie Commission and NCES have had to be revised downward several times because education is only conditioned to growth and not to contraction or even stabilization. Like Froomkin, Dresch strongly implies that the saturation of the job market for college graduates must be considered (Dresch, 1975).

Changing Patterns in Enrollments

Forecasting enrollments in the 1970's has become increasingly difficult for a variety of reasons. First, the average age of persons attending institutions of higher education has changed. In 1967, 36 percent of the students ranged in age from 18 to 19 years, 28 percent from 20 to 21 years, 16 percent from 22 to 24 years, and only 16 percent over 24 years of age. In 1977 the percentage of 18 to 19 year old students dropped to 25, 20 to 21 year olds to 21 percent, 22 to 24 year olds remained 16 percent, and 35 percent were now over 24 years old (Census Bureau, 1978). In the years ahead, the number of 18 to 19 year olds will decrease even more because the birth rate has decreased and, therefore, fewer people will be in the 18 to 19 year old cohort. Concomitant to the age increase has been a change in the mix of full and part-time students. Since the older students usually are employed and married, they are less likely to attend college full time. In 1963, 67 percent of students were full time; in 1976 the percentage had dropped to 60. NCES (1978) predicts that in 1986, 48 to 55 percent of all students will be full time. In the 2-year colleges in 1963, the percentage of full-time students was 50; in 1976 the percentage had dropped to 42, and the estimates for 1986 by NCES (1978) is 34 to 39 percent.

In the late 1960's and early 1970's higher education moved toward universal access to higher education (Carnegie, 1971). As a result of federal legislation and federal funding, more minority and women students now are attending college than ever before. In 1963, 38 percent of all students were women; in 1976, 47 percent were women, and NCES (1978) estimates that it could be as high as 50 percent by 1986. In 1963 in the 2-year colleges, 37 percent of the students were women and in 1976, 49 percent were women; NCES estimates that women could comprise 54 percent by 1986.

Similar trends have been observed in the enrollment of blacks. From 1964 to 1970 the number of blacks increased from 234,000 to 522,000 representing an increase of 123 percent for all institutions of higher education (Carnegie, 1971). By 1976 the total black enrollment rose to over 1,000,000, accounting for 9 percent of the total college enrollment (SREB, 1977 and 1978). As the enrollment of blacks increased, the percentage of blacks attending predominately black colleges decreased from 60 percent in 1972 to 50 percent in 1976. Fifty percent of all black students enrolled in college went to the community colleges (Monroe, 1973). Thirty-one percent of the students enrolled in community colleges in 1972 were black (Census, 1973). In addition, Glenny (1980) points out that the birth rate of minorities during the past decade has been significantly higher than for the mature white population. This will result in a different ethnic mix in the college age cohort and, because attendance patterns vary among ethnic groups, will, thereby, affect both the enrollment size and the programs and courses selected.

Through increased state and federal aid, more students from low-income families have been encouraged to attend college. Recent legislation requiring institutions to accept handicapped students should lead to an increase from this group. Since increased numbers of nontraditional students are attending institutions of higher education, traditional methods of forecasting enrollments were inadequate because these students have different attendance patterns which were not reflected in most previous forecasts.

While the growth rate in college enrollments is declining because of the decrease in the normal college age cohort, institutions of higher education have begun searching for new types of students. To attract such students, new types of curricula have been developed. More occupational and vocational programs are now being incorporated into the curriculum and remedial work is being offered by many colleges to attract students with academic deficiencies. Through federal legislation, the U.S. government is encouraging higher education to open its doors for "universal access" (Carnegie, 1973).

One critical factor seems to have been ignored in enrollment projections to date. Glenny (1980), among others, points up the tremendous growth of postsecondary educational programs sponsored by other than colleges and universities. This is especially true in the noncredit, nondegree programs. Not only is the total potential student population decreasing, but the available cohort must be more widely shared.

Kibbee (1973), in a paper given to the Association of Institutional Research, claims that higher education has not recognized the existence of certain basic principles in a democratic society. First, excluding people from an opportunity in our society, such as education, is inversely proportional to the percentage of the population that is included and to the importance of that opportunity for participation in future benefits of the society. Second, excluding people from a benefit of our society is inversely proportional to the public tax dollars invested in providing that benefit. Expenditures in higher education exceed $30,000,000,000 annually, roughly two-thirds of which are tax dollars. Higher education is now "big business". When an industry has grown as large and as fast as higher education, it can hardly go unnoticed and the demand for it increases.

Daniel Bell (1968) arrives at the same conclusions as Kibbee, but bases his conclusions on "Tocqueville's Law", which says, "in a society pledged to the idea of equality, what the few have today, the many will demand tomorrow." Higher education is an example of "Tocqueville's Law" in operation. The disadvantaged and those people who for various reasons have not had the opportunity to attend institutions of higher education have demanded the opportunity. To meet the postsecondary needs of all people who are capable and who desire the opportunity to attend institutions of higher education, the community college was established. Community colleges are meeting the needs of people with a wide variety of backgrounds and abilities (through a wide variety of programs). This diversity in the community college population makes forecasting enrollment in these institutions very difficult indeed.

Enrollment patterns are changing because the type of students attending institutions of higher education is changing. The number of adults attending colleges and universities, for example, is increasing. The Carnegie Commission in 1973 reported that 42 percent of all students on college and university campuses were adults. Since adult students tend to be part-time students taking both degree-credit courses and nondegee-credit courses, the number of students enrolled in full-time degree-credit programs is predicted to decline (Carnegie, 1973). Full-time degree-credit courses are usually taken by younger students who come directly from high school. Because adult enrollment is increasing, the Carnegie Commission (1973) predicts that by the year 2000 between 50 and 54 percent of the population will have 1 or more years of college. The percentage of people having 4 or more years of college by the year 2000 is estimated to be between 30 and 34 percent. At present only about 20 percent have 4 or more years of college. If "Tocqueville's Law" holds, as more people attend college, more will demand the right to attend.

National enrollment trends, however, have not always been indicative of state and institutional trends (Carnegie, 1973). Between 1970 and 1972, for example, national enrollment rates increased by 8 percent while several state enrollments increased by more than twice the national average (Nevada and South Carolina, 33 percent; Delaware, 19 percent; and Vermont, Virginia, and Wyoming, 18 percent). In other states, such as Minnesota, Montana, New Hampshire, and North and South Dakota, the enrollment rate significantly decreased. Demographic factors such as total population were found to have little relationship with state and institutional growth rates (Carnegie, 1973). Since national enrollment patterns differ from state and institutional enrollments, the same assumptions may not be relevant to states and institutions. At the 1976 forum of the Association of Institutional Research, Norris (1976) emphasized the importance of each institution developing its own formula for projecting enrollments because each institution differs in such characteristics as goals and missions, types of students served, and admission policies. Since state enrollment trends differ from national enrollment trends, states likewise may need to consider basing their enrollment projections on different assumptions, and, therefore, different factors. In the next section relevant factors for national, state, and institutional projections will be discussed.

Relevant Factors to Predicting Enrollments

The basic assumptions underlying each enrollment projection determine the type of factors used in the projection. Factors found relevant to projecting enrollments can be grouped into three general categories: demographic, economic, and institutional and are listed in Tables 2.1, 2.2, and 2.3, respectively. Demographic factors have been used in enrollment projections more often than any other category of variables. Several researchers, such as Leslie and Miller, Froomkin, Freeman, and Dresch, have suggested economic factors as alternatives to using demographic factors since recent enrollment projections using demographic factors have failed to produce accurate projections. Institutional factors are those factors directly related to the institution. Many institutional factors can only be subjectively assessed, and, hence, are not quantifiable. Institutional factors, such as those suggested by Bowen, are listed in Table 2.3. Many demographic variables could also be classified as institutional factors if they are particularized to an institution.

The usefulness of many of these factors by institutions are sharply limited by the validity of available data, or by the subjective nature of the factor. In addition certain variables are only relevant to certain types of institutions; e.g., first professional and graduate enrollment are relevant only to institutions of higher education with graduate or professional schools. At this time statistics on the draft are not relevant. As noted earlier by Folger (1974), the data base for institutions of higher education needs to be expanded if institutions, states, and national forecasters are to increase their accuracy.

TABLE 2.1 - DEMOGRAPHIC FACTORS

Factor	Researchers
1) College going rate	Carnegie, Commonwealth of Va.
2) Population	Carnegie, Cartter, U.S. Census Bureau
3) Age distribution	Carnegie, Froomkin, Newton, NCES
4) Sex	Carnegie
5) High school graduates	Wasik, Carnegie, Froomkin, Newton
6) Birth rates	Carnegie, U.S. Census Bureau
7) Racial composition	Carnegie, Newton
8) Degree-credit enrollment	Carnegie, Hollander
9) Part-time enrollment	NCES, Carnegie
10) Full-time enrollment	NCES, Carnegie

TABLE 2.1 - DEMOGRAPHIC FACTORS (Concluded)

Factor	Researchers
11) Number and type of institution	Bowen
12) College attendance rates	Froomkin, Bowen, Cartter
13) Graduate enrollment	Carnegie
14) First professional enrollment	Carnegie
15) Nondegree-credit enrollment	Carnegie
16) Pupil-to-teacher ratios	Carnegie
17) Draft	Conner, Carnegie, Wasik
18) Number of college graduates	Froomkin
19) Education characteristics of active adult population (AAP)	Dresch
20) Education characteristics of persons leaving the AAP	Dresch
21) Participation rates	Newton
22) Previous schooling	Russell, Hoffman
23) Educational background of parents	Russell, Hoffman
24) Public school enrollment	Conner
25) Persons below public school age	Conner
26) Persons enrolled in other institutions	Conner
27) Persons having a degree	Conner

TABLE 2.2 - ECONOMIC FACTORS

Factors	Researchers
1) Number of jobs available	Froomkin, Freeman
2) Type of jobs available	Froomkin
3) Growth of economy	Leslie and Miller
4) Relative educated wage	Dresch
5) National economic indicators	Wasik, Salley
6) Jobs in service area	Bowen
7) Local unemployment rate	Salley

TABLE 2.3 - INSTITUTIONAL FACTORS

Factors	Researchers
1) Relevancy and attractiveness of program	Bowen, Conner
2) Convenience of times and places	Bowen
3) Admission requirements	Bowen, Conner
4) Tuition charges	Bowen
5) Terms of financial aid	Bowen
6) Release time from work	Bowen
7) New degrees	Leslie and Miller
8) Attrition rates	Lightfield
9) Seasonal factors	Salley
10) Plans of high school students	Russell and Hoffman

TABLE 2.3 - INSTITUTIONAL FACTORS (Concluded)

Factors	Researchers
11) Degree of influence of certain persons	Russell and Hoffman
12) Impact of future events	Russell and Hoffman
13) Student aspiration	Russell and Hoffman
14) Faculty size	Russell and Hoffman
15) Academic standing	Russell and Hoffman
16) Type of community	Russell and Hoffman
17) Area industry	Conner
18) Class level distributions in student flow models	HEEP, Lightfield, Wasik, Koenig, Thonstad, Dietze, Caspar, Gani

Certain categories of demographic data have had low priority. In the past decade the goals of higher education have been broadened to include new types of students. The era of "universal access" to higher education has made the planing of higher education more complex than in previous decades. The number of nontraditional students are increasing while the number of traditional college students are decreasing. To meet the needs for more precise enrollment projections, data bases must be broadened to include the changing growth patterns of higher education.

Accuracy of Enrollment Projections

The accuracy of enrollment projections can be judged only when data become available. For the forecasts of the 60's such data are now available. In 1961 Ronald Thompson under projected 1970 enrollments by 5 percent; the U.S. Census Bureau under projected 1970 enrollments by 5 percent for their high projection and 29 percent for their low projection; and the U.S. Office of Education under projected bachelor degrees awarded by 13 percent, master degrees by 29 percent, and doctor degrees by 33 percent (Folger, 1974). This would appear to point up the inaccuracies of past projections and the need for more accurate projection techniques.

The enrollment projections made in the 1970's for the 1980's cannot be judged by comparison with actual data because the data are not yet available. However, certain statistical analyses such as the mean square error, the average error, and multiple coefficient of determination are useful for evaluating the accuracy of predictions. The multiple coefficient of determination cannot, however, be obtained for all forecasting designs, and is, therefore, only useful where applicable. Since none of the enrollment projection studies reviewed reported these statistical quantities, judgment of the accuracy of the forecasts is difficult.

Methods of Forecasting

The three generally used methods of statistical forecasting are intuition, causal, and extrapolation. The classical or traditional approach is the intuitive method. The extrapolation method, on the other hand, has been the method most often used recently since most institutions have assumed a continual growth pattern. The causal methods are based on the scientific method. Each of these methods involve many specific forecasting techniques, some of which will be discussed in this section. Each technique has advantages and disadvantages in different situations. Since the purpose of this study is to identify those factors having the greatest influence on enrollments in a comprehensive urban community college and to utilize these factors in a forecasting model, each technique will be examined thoroughly as to data requirements, specific assumptions, and time requirements to determine its appropriateness.

The Intuitive Method

The intuitive method is a judgmental approach that covers the whole gamut of nonexplicit inferences. It was the most extensively used method before the massive accumulation of data was made possible by the development of the computer. "Guesstimation" and intuitive hunches were the most commonly used techniques. An expert's opinion or an individual's feeling about a situation was the sole basis upon which forecasts were made. Since the logic behind such approaches cannot be explicitly stated, no basis existed for evaluating why a forecast was wrong and rationally improving forecasts was nearly impossible. When the judgmental approach, unaided by explicit procedures is used now, its advocates rationalize that the human mind is the most sensitive and comprehensive evaluator of complex and diverse evidence on what the future holds (Ascher, 1978). When the logic behind the judgmental approach can be explicitly stated, its conclusions are more acceptable because less chance exists for bias such as political or sociological prejudices. Whether the conclusions are any more reliable can only be determined over time. If the forecasts are merely guesses or hunches, the reliability of the results are minimal. When a logical set of procedures is used in the judgment forecast, the accuracy increases. To increase the forecasts made by one expert or individual, the single expert could be replaced by a group of experts who arrive at a group consensus. The Delphi Method is such an approach. An alternative approach for forecasting college enrollments that would yield more reliable results might be a survey of the potential clientele.

The intuitive approach is infrequently employed as the only method to forecast enrollments. In fact no cases were discovered where the researcher admitted to only using this method. It was, however, almost always used to some degree in such studies. For example, when variables are selected for a causal model, the factors are selected by intuition. Montgomery College in Maryland, for example, even incorporates judgment as one of the three essentials of its model (O.I.R., Montgomery College, 1977). Although most enrollment forecasting studies are classified as either causal or extrapolation approaches, intuition always plays a role. Froomkin and Bowen, as previously discussed, are examples of the intuitive approach to forecasting.

The Causal Methods

The causal methods of statistical forecasting attempt to forecast effects by determining their causes. The causes may be economic, political, sociological or demographic in nature and may be unpredictable, or the time lag between the cause and effect may be so short that the information is useless for predictive purposes. In most instances mathematical and/or statistical descriptions are used to describe the relationships between the cause and effects. Only those causes that can be quantitatively determined at a long enough time period prior to the effect occurring can really be helpful in forecasting. Techniques classified as causal methods are cohort-survival, ratio, multiple correlation and regression, path analytical, and systems of equations (Wing, 1974).

Cohort-Survival Techniques

In the cohort-survival technique, the cohort is defined as "a group of individuals having some common classification trait or traits" (Wing, 1974). A cohort for the community college, for example, would be those young people who come directly from high school; another would be those students who return immediately following a quarter in which they had been enrolled. This technique is based on the assumptions that net migration, mortality, the percentage of high school graduates going directly to the local community college, and the percentage of returning students remain constant over time. If these assumptions are not valid, adjustments must be made. If, for example, the percentage of high school graduates attending the local community college is increasing, other causal or extrapolation forecasting techniques must also be used.

The cohort-survival technique has been used successfully to predict enrollment in the public schools from kindergarten through twelfth grade where general patterns of grade repeating and the grade progression ratios (the fraction of students in one grade level that continue to the next grade level in the next year) are stable over time. This technique could also be extended to those first-time students at the local community college who come directly from high school.

To use this technique the first step would be the calculation of the grade progression ratios for each grade used in the progression. If only grades 9 through 12 were used for the calculation, the general assumptions would have a higher probability of being stable over time since the time lag would be much shorter. A second consideration might be the selection of each high school within the service area of the community collge versus using the total high schools within each political subdivision. Each high school could be used individually or the sum total within each political district could be used. The final steps would include the calculation of the high school graduation rate and the college going rate (penetration ratio) for the local community college. A causal relationship is assumed to exist between the number of high school graduates and the number of first-time students enrolled in college. Since all high schools must now keep records of the number of graduates, the data needed are readily accessible. With the availability of computers, the calculations required in this forecasting technique can be done in very short periods of time. The major disadvantage of forecasting enrollments by this technique is that it is useful only for a small segment of the enrollees at a college because no known cohort exists for the adult students.

At the University of California, Berkeley campus, the cohort-survival technique has proved beneficial in helping the university prepare for the rushed period at the opening of the fall quarter as well as for long range (5 to 10-year) enrollment forecasts (Suslow, 1977). The cohort-survival method uses class level at the time of admission. A survival or persistence rate is obtained for each student from fall term through each succeeding term until he either graduates or leaves the university. The persistence rate is determined by the overall behavior of the entire group that constituted the original cohort for the specific term being considered.

When a residual of current enrollments was traced to cohorts entering 10, 20, or 30 years in the past, the first impulse was to reject the possibility of their having a significant impact on a projection concerned with nearly 20,000 graduates. The model's validity, however, rests on its ability to account for all sources of current enrollments. When other cohorts (no matter how small their contribution) are discovered, they are included in the model. Probabilities of survival are then established and are applied to future incoming classes. Persistence rates, however, must be checked periodically to maintain the model's validity over time. For example, when the fall 1955 and fall 1960 cohorts were compared to the fall 1979 through 1974 freshman cohorts, the latter group had higher persistence rates. When the model explored all relevant combinations--fall to winter, winter to spring, and so on--the model revealed that 19 percent were freshmen 1 year later, 64 percent were sophomores, 1 percent juniors, and one-tenth percent seniors. Since the jumps in 1 year from freshman to junior or senior level are obviously not accurate measurements of student behavior, the model revealed administrative and system problems in making accurate and timely measurements of student records or changes in those records.

One of the major problems found with the model at Berkeley is that opening fall enrollment data cannot be obtained early enough to incorporate them in the next round of forecasts. Other methods are, therefore, used as ancillary support for the model. The benefits of the model are twofold: first, it depends on measurements of student academic behavior rather than official expectations of how undergraduates move through their curricula toward a degree; and secondly, the original observation and subsequent confirmations that student survival rates established over time are more consistent than the term-by-term variations in class standing yielded by administrative registration processes (Suslow, 1977).

The Ratio Techniques

Ratio techniques are based on the existence of a relationship between a particular ratio such as grade progression ratios and some effect such as a particular segment of the enrollment produced by it. The cohort-survival technique is a special case of the ratio technique that has been applied extensively by the public schools and in some cases very effectively applied to college enrollments. To determine if a relationship exists, simple correlations are computed (Wing, 1974). When a relationship is established, other forecasting techniques such as multiple correlation and regression or extrapolation may be needed in the final predictive process. If the ratio should be stable over time, it could be used as an entity in itself, as a predictor of enrollments, or as one of several predictors in a more complex model. Should only one ratio be needed to obtain a good forecast, the advantages are obvious: simplicity, minimal time needed for the calculations, and a minimal amount of data collection. Since most ratios in recent years have not been stable over time and the relationships may prove fortuitous, the method is only as powerful as the temporal variation of the relationship between it and the effect for which it is a predictor. In the cohort-survival technique the persistence or survival rate in the public schools was highly correlated with the enrollment the following year (Wing, 1974). Ratios found to have causal relationships with college enrollments are persistence rates, participation rates, age distribution ratios, attrition rates, and graduation rates. In most instances, however, these ratios are found to be most effective in forecasting enrollments when combined with other factors. Newton (1976), for example, in a projection model for the institutions of higher education in the State of Pennsylvania computed the participation rates for the two sexes, two racial categories, and five different age groups to arrive at a forecast.

Multiple Correlation and Regression Techniques

Since a wide variety of causal factors can be examined using multiple correlation and regression techniques, it is probably the most powerful and most widely used method of statistical forecasting today. In forecasting community college enrollments, for example, the probability of finding one factor that will accurately predict enrollments is small since the population is so diverse. Most of the first-time student enrollees do not come directly from high school, but are older students who are employed and enroll for a variety of reasons. In the regression technique, an attempt is made to determine the relationship between a dependent variable (e.g., enrollments) and one or more variables sometimes called predictors or regressor variables. The general form of the regression model is $y = f(x)$, where $f(x)$ is a function to be determined. These functional relationships may be linear, exponential, logarithmic, harmonic, or a combination of these forms. The basic assumption underlying the model is that there is independence in the sample taken for the variables and the underlying data can be represented by trends (Lovell, 1971). In most statistical problems where regression techniques are used, the basic requirement is the existence of a "good fit" between the function and the past values of the dependent variable. In statistical forecasting, on the other hand, the criteria are not the various measures of fit described in most textbooks dealing with regression, but criterion such as the mean square forecasting error that refers directly to the forecasting use made of the regression model (Gilchrist, 1976).

Choosing the number of variables and the likely candidates for the dependent variables are critical to using multiple correlation and regression techniques for forecasting. Whether a model is global or local strongly affects the selection of the variables. In a global model the structure is assumed to be stable for a long time period and, therefore, will be the truth about the underlying structure. A local model, on the other hand, is one which may be true for the short run, but not for the long run. If a global model is chosen, the variables can be chosen on the following basis (Gilchrist, 1976):

1. By repeated examinations of the correlation between the fitted regression equation and the dependent variable
2. By examining the changes produced when variables are added and removed from the model
3. By examining the significance of the various regression coefficients

Selection of regressor variables are, therefore, based on such statistics as the correlation coefficient and the "t" statistic (Gilchrist, 1976).

In most instances a local model is more useful for forecasting purposes because experience shows that the more recent the data, the greater the probability of a good forecast. Variables which provide the best explanation of the most recent data should therefore be sought, rather than historical data which may no longer be relevant. When only recent data are used, the number of observations available is usually small which limits the number of variables that can be used in a regression model. When the number of variables exceeds the number of observations, the variances in the coefficients become so large that the regression model becomes unreliable for forecasting purposes.

When a local model is used, the basic assumptions underlying a global model will not be valid. The correlation coefficient and "t" statistic, therefore, must be replaced by modified versions based on discounted or weighted statistics. Types of variables that have been used successfully in local model forecasting situations are: (Gilchrist, 1976).

1. Lagged time variables; the influence of a particular variable may occur some time prior to its actual effect taking place. For example, the present cost of living index may well influence next years' enrollment.
2. Economic variables; if an economic variable is derived at the government level, these values may be actual values, forecast values, or lagged time values. There are, of course, many pitfalls in using economic variables.
3. Variables that may react differently with the dependent variable. For example, one independent variable might vary

linearly while another varies exponentially with the dependent variable.

Regression models may either be stochastic or deterministic. A deterministic model is one that has a perfect mathematical form with no chance or random element to influence it; a stochastic model, on the other hand, has chance elements which must also be predicted. A deterministic model rarely describes the real world, but most models have a deterministic element. Autoregression models by their very nature are stochastic. In autoregression a variable is regressed against itself at some prior time with the regressor variable usually given some weighting factor. The weighting factor is some function of the chance element and a given parameter. The chance element is the total effect of the unpredictable factors in the situation at some time "t". Since the chance element must also be predicted, the process is stochastic (Gilchrist, 1976). In projecting community college enrollments, the fall enrollment might well depend on students returning from the summer, spring, winter, and previous fall quarters. By regressing fall enrollments against previous enrollments, a relationship may be found to exist with the chance element being the drop-ins and drop-outs.

As previously noted, multiple correlation and regression techniques are the most commonly used techniques today, especially in attempting to understand student behavior patterns; i.e. student demand estimation. Without understanding the sources of the demand for admission by potential students, Wing (1974) believes that the chances of being able to develop reliable forecasts in a period of change such as has existed in the 1970's are greatly diminished.

Wasik (1971) at North Carolina State University suggests two models, one of which is a regression model for estimating the numbers of students to be served by the community college. The independent variables suggested by Wasik for inclusion in his model are: numbers of high school graduates for 2 preceding years, required local draft board needs for the 2 preceding years (this variable may or may not apply), estimate of economic activity, and county population. Wasik does not show an actual application of his model and its validity is difficult to assess.

The National Center for Education Statistics (NCES, 1978) uses simple linear regression for most of its national enrollment projections. Straight lines are fitted to a ratio such as enrollment rates as the dependent variable and time in years as the independent variable unless straight line growth appears unrealistic. When growth rates appear exponential, for instance, a logarithmic growth curve is used.

A number of other regression models are currently in use, Wayne Smith (1970) of the Office of Advanced Planning at UCLA has been using regression models in connection with other techniques to develop estimates of numbers of students. The Renssalaer Model developed by the Renssalaer Research Corporation (1970) couples regression techniques with a Markovian process to obtain enrollments. Banks (1970) at the University of Georgia uses three models, one of which is a simple regression model and a second is a multiple regression model to predict higher educational enrollments in the State of Georgia.

Path Analytical Models

Path-analytical models are extensions of multiple correlation and regression models. The relationship between the dependent variable and the chosen independent variables must be established prior to their use in the regression equation. Its usefulness is that it is a method of establishing the factors as significant before their use in the forecasting model. Wing (1974) and Graziano (1972) in the State of Illinois, Tsai (1973) in a general study, and Bradley and Lehmann (1975) for Empire State College are a few examples of studies where factors that limit growth in enrollment have been isolated. Some of these factors very easily could be used in projection studies as indicators for changing enrollment patterns especially in specific situations. Tsai did use the sociological determinants of college attendance which he found relevant in several causal models.

Sets of Equations

Finally, sets of equations may be derivied to establish a model that has as its basis a causal relationship. If the sets of equations only rely on historical data, they are considered extrapolation techniques. If, on the other hand, a causal relationship exists, they are causal techniques. Two models which are currently being used in forecasting enrollments and which could be considered systems of equations are the student flow models and the Markov transition models, either of which could be themselves classified as causal or extrapolation techniques, depending on the underlying assumptions.

A Markov model is simply a formal mathematical structure which consists of transformations from one state to another during an increment of time. The basic assumptions underlying Markov transition matrices are that the transition must depend only on the present state and be independent of the past. A second assumption is that the system be stationary. A system may be considered stationary if: (1) The time interval chosen is small or (2) a change takes place over a long time period. Since changes in the educational system usually take place over long time periods, the system can be considered stationary, thus obeying the second Markov assumption.

Student flow models may be considered Markov transitions if, and only if, the two basic assumptions are obeyed. When student flow models use Markov transfers to forecast enrollments, probabilities are calculated for the movement between two levels, such as the probable number of freshmen who will become sophomores or the probable number of students who will return in the winter quarter following completion of the fall quarter. Since the probabilities may not remain stable over time, they should be recalculated periodically. Other sets of equations may be developed to describe interrelationships among enrollments and other external and internal factors. When too many equations are included in a model, however, the explanation of the model to high authorities may become difficult.

Examples of student flow models used for forecasting purposes are numerous. Lightfield (1975) at Mercer County Community College developed a tracking system in order to study attrition which developed into a multidimensional system, one of its outcomes being enrollment forecasts. Wasik (1971) at North Carolina State University developed a flow model based on Markov transformations for projecting community college enrollments. Like his regression model previously discussed, he shows no application of the model, and, therefore, the validity of the model is questionable. The Office of Program Planning and Fiscal Management in the State of Washington (1970) has developed a student flow model which uses the Markovian transformations to project undergraduate enrollment. The Higher Education Enrollment Projection Model (HEEP) is an interinstitutional as well as intrainstitutional model. One student flow model developed for a state system is the Michigan State Model developed by Koenig et al (1968). Foreign researchers such as Thonstad (1967) in Norway, Dietze (1969) and Caspar (1969) in Germany, and Gani (1963) in Australia have developed student flow models using Markov theory as long as 15 years ago. All student flow models are not developed for the primary purpose of predicting enrollments, however. In most cases enrollment projecting is a secondary outcome. The primary purpose of most student flow models is for planning purposes since identifying the way students move through an institution may reveal a great deal of useful information.

Extrapolation Methods

Extrapolation methods are based on the continuance of past trends. An implicit assumption exists that the past is indicative of the future. The main advantage of extrapolation methods is that the only input data required is historical. These methods are generally used in two situations: when trends in the past will, in fact, continue in the future and when too little is known about relationships to permit the development of causal models (Wing, 1974). Extrapolation methods include several curve fitting techniques: the global constant mean model, the local constant mean model (sometimes called the moving averages model), the exponential smoothing model, the linear trend model, the polynomial model, the exponential model, and spectral analysis. Curve fitting models may even be used to forecast parameters for causal models. In all these models, the researcher is looking for trends that appear to fit specific curves. Although many of these techniques are similar to causal techniques, the difference is the rationale behind the models. In curve fitting models no rationale exists for their use while, in the causal models the basic assumption is that a recognizable cause exists which produces an effect.

Constant Mean Model

In the constant mean model, the mean of past data is the forecast for the next time period. The constant mean model may be expressed in three forms: the explicit form, the recurrence form, and the error correction form. Where all the data is explicitly included in the mean, the explicit form is used. The new forecast is simply the mean of all the data--old and new. When data are included on a continuing basis, the recurrence form is used. The new forecast can be obtained from the old forecast and the newly obtained data by simply adding the new observation to a multiple of the old forecast and dividing by the total number of observations to obtain a new mean. In the error correction form, a multiple of the forecast error (the difference between the forecast and the actual observation when obtained) is added to the last forecast to obtain the new forecast. The multiple used is one divided by the total number of observations. As the number of forecasts increases, the new forecast becomes less sensitive to forecasting errors. A large change in structure would produce large errors, but the forecast change would be small due to the large number of observations. The averaging operation, therefore, has the effect of reducing the random variation and leaving an estimate of the mean. Since the averaging takes place over all the data, the model is considered global. For a global model in any form, the variable to be forecasted (e.g., enrollments) must be constant over time. Since a trend such as enrollments is unlikely to be constant, the use of a global mean model for forecasting is generally not considered to be a good choice. Understanding the basic concepts underlying this model are important since these concepts are used in the next two models to be discussed--the local constant mean model and the exponential smoothing model (Gilchrist, 1976).

The structure for the global constant mean model and the local constant mean model are the same and no real empirical differences exist although the basic assumption of the two models differs from the local model. In the local model, the mean need only be constant for the time period over which the average is taken. A local model, therefore, tends to be a better basis for forecasting purposes than a global model because it responds to structural changes. In the local constant mean model, the data is divided into successive groups. Each successive group after the first group adds the most recent observation until all observations are used. When the mean of each group is determined, it can only be regarded as an estimate of the mean at the center of its group. Since these averages move from the earliest to the most recent data, they are often referred to as "moving averages" (Wing, 1974). A recurrence form can be used if new observations are obtained. Only when the moving averages are constant or only varying slightly can this model be used for forecasting purposes. Its most practical use is to smooth data or manipulate data. At the option of the researcher, the number of historical points to be included in the averages or the number of averages to be considered can be changed. A moving average may show, therefore, where a structural change occurs, and when a new model or technique is needed (Gilchrist, 1976).

If the most recent data used in a local constant mean model is weighted, the technique is much more useful for the purpose of forecasting. When data are weighted in statistics, the technique is called discounting or discounted statistics. Although there are many ways in which the moving averages can be weighted, the most appropriate method and the one most often used is an exponential set of weights. Since forecasting is usually a continuing process where new data are continually added, the recurrence form of exponential weighting appear to be the best form to use (Gilchrist, 1976). When this form is used, the equation is

New Forecast = [(1 - a) x New Observation] + [a x Old Forecast]

where a is a constant whose value ranges from $0<a\leq 1$. Since a is a forecasting parameter as distinct from a parameter of the model, it is usually chosen so as to minimize the mean square error of forecasting over a trial period. When a is small, only the more recent data influences the forecast. As a approaches 1, earlier data begin to influence the forecast. At $a = 1$, the older data have the most impact on the forecast. If the forecast error is included in the recurrence equation, the equation becomes more sensitive to changes in the data. The new forecast can be obtained from the old forecast by adding a fraction $1 - a$ of the last error made in the forecast. In the global constant mean model, it will be remembered, the error term divided by the number

of observations was added to the old forecast to obtain the new one. As the number of observations becomes large, the forecast becomes less sensitive to forecast errors. Since a is independent of the number of observations, the error correction form of the exponential smoothing technique is much more sensitive to forecast errors and, therefore, to changes that occur in the population. Unless the researcher is certain that a constant mean model has applied in the past and will continue to apply in the future, exponential smoothing is safer to use than the ordinary mean (Gilchrist, 1976).

Polynomial Models

In polynomial models, the assumption underlying the first three models discussed is relaxed; i.e., that the dependent variable (in this case, enrollments) will be constant over time. A restriction common to all polynomial models is that there must be at least as many historical data points available as there are parameters to be estimated (Wing, 1974). In the first order polynomial, the linear trend model, historical data is matched to a straight line based on the method of least squares. The general form of the linear trend model is

$$x_o = a + bt + e_i$$

where a and b are parameters to be estimated, t represents time, e_i is a sequence of independent random variables with expectation equal to 0 and x is the new forecast. The method is essentially the same as the linear regression model except that no causal relationship is assumed to exist. When $e_i \neq 0$ the model is biased and the process becomes stochastic. At $t = 0$, the linear trend model reduces to the constant mean model; i.e., a is simply the mean in the constant mean model (Gilchrist, 1976). The general form for higher order polynomials is

$$x_0 = a + bt + ct^2 + \ldots\ldots\ldots + e_i$$

where a, b, and c, etc. are again parameters to be estimated from historical data by the method of least squares. A higher order polynomial should only be used when the data show *strong* indications of continuing along a particular curve. In many cases the data trend may change overnight, making the model useless. For enrollment trends, Wing (1974) believes higher order polynomials have little application as a forecasting method for enrollments. The technique for solving higher order polynomials is the same as for multiple regression.

Nonlinear Growth Models

Logarithmic, exponential, Gompertz, and logistic models are examples of nonlinear growth models of which the exponential is the most important. Many examples of exponential growth patterns can be found in economics and science. Essentially, the exponential model is a special case of the logarithmic model to the base e. This class of models differs from polynomical models in that they show regions of very rapid change and of maximum (or minimum) values (Gilchrist, 1976). Since more examples of the exponential model exist and the characteristics of this whole class of models are similar, only the exponential model will be discussed in detail.

Exponential models are nonlinear models that assume the time rate of growth (or decay) of a forecast is directly proportional to the value of the forecast variable itself. Because enrollment growth has varied in the 1970's, the actual enrollment trend has become controversial. According to Wing (1974), exponential models have little value for purposes of forecasting enrollments because the exponential growth curve is one that continues to grow with greater and greater rapidity, and the growth rate of enrollments in the 1970's has leveled off and even declined. Leslie and Miller (1974), on the other hand, see the growth in enrollments tied to the economy which is theorized by many economists as growing exponentially. The growth pattern of enrollments in the 1970's is just a minor perturbation in the total growth pattern, and the growth in enrollments will very shortly resume an exponential growth pattern. The general form of an exponential model is

$$x_t = ae^{bt} + e_i$$

where a and b are parameters to be estimated, t represents time, e_i is the random error which is assumed to be 0 in a deterministic model and $\neq 0$ in a stochastic model and x_t is the forecast at some time t.

The exponential model can be very good at producing short term forecasts, but the accuracy of longer term forecasts is questionable because an element of doubt must exist that any real situation will continue to increase indefinitely with greater and greater rapidity (Gilchrist, 1976). In the initial stages of development, a new product which competes with other similar products, will often expand its market from 0 to some value until it acquires a certain percentage of the market. This initial growth period may follow some exponential growth model until it achieves a given percentage of the market (Gilchrist, 1976).

When investigating enrollment patterns, the exponential model should be considered at least when the system is in its initial growth period. Since the community college movement is relatively young in its development, the possibility exists that their enrollments may follow an exponential growth curve. Individual institutions, when first established may, likewise, follow an exponential growth pattern in their enrollments. The influx of students is greatest in the first years after the establishment of a community college and levels off as the initial demand is satisfied. Individual segments of the college population also might continue along an exponential curve. The number of part-time students, for example, is increasing at a rapid rate and may well follow an exponential growth curve.

Spectral Analysis

Spectral analysis involves the use of harmonic functions (sine and cosine functions) in the model. In spectral analysis a large number of data points are fitted to an infinite series of sines and cosine functions. Since the model requires at least 25 historical data points (Wing, 1974), the possibility of using spectral analysis is limited due to the lack of sufficient historical data. Prior to 1967-68, Jellema (1972) reported that 13 percent of all institutions had no data on freshman applicants or the number admitted; 5 percent could not even supply headcount enrollment. Also, when 25 data points are required, the data becomes too far removed to be useful in forecasting (a second disadvantage to spectral analysis).

The Time Series Decomposition Model

The time series decomposition model is an extrapolation model which assumes that some pattern or combination of patterns are recurring over time. Two factors in this model are important: the series to be forecast and the period of time under study (Wheelwright and Makridalus, 1977). In most instances the basic time series contains subparts which must be identified to obtain an accurate forecast. This situation occurs when a seasonal pattern is present. The basic pattern must then be divided into its subparts and is often broken into four parts: the trend factor, the cyclic factor, the seasonal factor, and the error factor (Wheelwright and Makridalus, 1977). These factors may be additive or multiplicative; i.e.,

$$F = T \times C \times I \times S \text{ or}$$

$$F = T + C + I + S$$

where F is the forecast, T is the trend pattern, C is the cyclic pattern, I is the irregular or error term, and S is the seasonal index. Patterns may exist, however, when one or more of these factors are not present. The trend factor is usually the time series portion of the basic pattern; the cyclic term is present when the pattern shows periodic upward and downward swings. It is used to describe movements which are slower and much less predictable than the seasonal component (Gilchrist, 1976), such as economic or business trends. The seasonal factor is present when a repetitive nature exists over some short period of time (a month or quarter). To remove the seasonal variation, a moving average is used to obtain a seasonal index which is then used to adjust for the seasonal variation. Salley (1979) used this approach in forecasting enrollments at Georgia State University. He found that a time trend, a cyclic factor, and seasonal factor existed in the data. Credit hours per quarter were seasonalized using the moving average technique. When the seasonal variation and the basic time trend were removed, a cyclic pattern emerged that correlated highly with two NBER indices. These indices were composites of leading and coincident economic indicators of the business cycle and are published in the Business Condition Digest monthly.

Enrollment Projections Employing Extrapolation Methods

Because higher education has historically been a growth industry, most institutions of higher education and their national agencies have forecast enrollments by extrapolating past enrollment trends upward (Bowen, 1974). Traditionally, enrollments have been considered a function of history. Enrollment patterns of the 1950's determined the projections of the 1960's and enrollment patterns of the 1960's determined the projections of the 1970's. Trends in the 1950's underestimated the enrollments of the 1960's whereas the trends of the 1960's overestimated the enrollments of the 1970's (Mangelson, 1974). Many studies were inaccurate because they were based on incorrect assumptions where, for example, the college going age would continue to be 18 to 22 years or where the birth rates would continue to be 2.7 births per woman. The average college going age is increasing because more older students are attending college (Carnegie, 1973). Since 1960 the U.S. Bureau of Census has revised birth rates downward three times to a low of 1.7 births per woman (Mangelson, 1974). In the latest Census population projections, the fertility rate was assumed to be 2.1 (Carnegie, 1975). Many enrollment projections based on the U.S. Bureau of Census population projections, therefore, have had to be revised, and new methods employed. For national enrollment projections, the U.S. Bureau of Census now uses the cohort method, the ratio method, and surveys. NCES (1978) uses regression, double exponential smoothing, and ratio techniques as a basis for their forecasts. Extrapolation techniques, therefore, are still the major techniques used for national forecasts. Many states such as the Commonwealth of Virginia use techniques similar to the national agencies. Virginia bases its forecasts on the college going rate and state population projections for the 18 to 34 age group (SCHEV, 1974).

The trend demographic model by Hollander (Mangelson, 1974) is an interesting forecast used for predicting full-time graduates in the State of New York. He predicts a 20 percent or greater decline in traditional undergraduate enrollments by 1990. He isolates the demographic effect on traditional undergraduate enrollments and applies regional variations.

Basis of Forecasting

In statistical forecasting, the basic requirement is the existence of a stable structure over some time period. The structure may be mathematical and/or statistical in nature. In science both stable statistical and mathematical structures exist. The stability of structure, however, does not require a precise law or relationship. If the chance element is constant over time, an equation or formula may be derived that has a certain probability of being correct at least part of the time. The smaller the chance element, the better the probability of having a reliable forecast. For example, all the molecules of air in a room, at least in principle, have a finite probability of leaving the room at the same time. That probability is, however, very, very small. Since the chance element is very small and stable over time, a statistical equation can be derived that predicts the number of molecules in the room at any given time with extreme accuracy. Thus, if an accurate forecast is to be achieved a stable structure must exist over time and any chance element must be small and stable over time.

Scientific Forecasting Procedures

Basic Steps

The scientific approach to forecasting usually consists of five basic steps: all data relevant to a particular situation are collected and reduced to a basic minimum (called data reduction), the most basic data are studied in depth for possible trends or relationships, a model (or models) is constructed from the structure implied by the in-depth study, and finally a forecast is made using the constructed model (Gilchrist, 1976).

Data are usually divided into two types: endogenous and exogenous. Endogenous data are internal to the system whose behavior is determined by the model itself; exogenous data are external and are not affected by the system. Exogenous data are predetermined by some other model or system. In many instances only endogenous data are needed to construct an accurate forecasting model. Endogenous data for an enrollment forecasting model might include first-time students, returning students, dropouts, or graduates. Exogenous data would include demographic, economic, political, and sociological factors. Selection of data can only be made on the basis of a thorough understanding of the situation. In fact, intuition very often plays a dominant role in data selection. If the selection is poor, the forecast will most likely be poor.

Data are reduced to a minimum for two basic reasons: identification of structure and reduction of error. Each variable has its own inherent inaccuracies. Increasing the number of variables must, therefore, increase the inaccuracy of the model. The criteria for data reduction are relevancy, reliability, and recency. Endogenous data more often satisfies these criteria as they directly affect the situation and are easier to obtain. When these data have been thoroughly studied and the behavior clearly understood, a mathematical or statistical model is the outcome. If no simple structure appears, more data may need to be included and studied. Exogenous data may now be considered for inclusion in the model or at least used to interpret the endogenous data as to causes of change (Gilchrist, 1976).

Identifying structures and building models can be very complicated and difficult. The most important step and first step after data reduction is the graphing of the data to investigate for simple trends. After a comprehensive, in-depth study of the data for trends is accomplished, the researcher should be able to formulate assumptions or hypotheses on which to build a model. The model is simply a mathematical or statistical description of the behavior of the data. The choice of a model is basic to the forecasting process, and often the most basic models are the only ones needed for forecasting. Sometimes, however, the need will arise to combine or modify them to a particular situation (Gilchrist, 1976).

Models have certain basic common characteristics. These general characteristics will now be considered. As noted previously, models may be global or local. In a global model the structure is highly stable and the chosen model is assumed to be true for all time. In a local model, the structure is assumed true only for the short run. Short term forecasting is usually considered to be <2 years and long term forecasting for more than 5 years (Ascher, 1978). For most forecasting purposes, the local model is probably preferable since such models will be more likely to pick up changes in trends more quickly. A local model is an approximation valid only locally in time. Some practical situations in which local models are used are:

1. Where not enough data are available to study the situation in detail;
2. Where, for reasons of simplicity, economy, or time a simpler model than may be indicated is used; and
3. Where the factors being forecast are known to be unstable, but the changes occur in a relatively slow fashion (Gilchrist, 1976).

When local models are used, the following methods are used in descending order: the intuitive method, the causal method and, finally, an indicator approach. The indicator approach is the use of preselected trends as leading signals of change in the situation. Extrapolation techniques are considered unattractive for use when the time span covered is short. They are, however, used as benchmarks in evaluating the performance of other short term techniques. Although extrapolation techniques may do the job better, practicing forecasters tend to use more complex techniques that incorporate more information, theory, and intuition (Ascher, 1978).

Local Forecasting Models

Numerous local models have been developed for various purposes. The major types have been described in a previous section; the discussion here will be limited to three models which have unique characteristics and have enjoyed some success in predicting enrollments for colleges.

In 1971, when enrollments dropped at the universities in the Province of Manitoba, the University Grant commission reexamined their methods of forecasting enrollments and arrived at a Postsecondary Demand and Enrollment Model which proved to be accurate within 2 percent (Russell, Hoffman, 1976). The model essentially surveys a sample of students in grades 10, 11, and 12 as to their plans after high school. The survey items include: the postsecondary plans of 10th, 11th, and 12th grade students; the degree of influence that individuals, parents, high school activities, and circumstances had on students' decisions; the impact of future events on students' decisions; the students' aspirations, expected career choices and long range plans; background questions concerning faculty size, previous schooling, and type of community where students grew up; students' academic standing; the students' choice of institutions and programs; optional questions concerning parents' education and family income; the financial situation in the province and the students' financial prospects; and reasons for students in grades 10 and 11 not completing high school. The model provides an intermediate range forecast for periods of 3 years which are, in turn, attached to traditional forecasting methods for years 4 through 15. One of the virtues of the model is its ability to show major shifts in student demand and it allows the institution to make adjustments before such shifts occur (Russell, Hoffman, 1976).

A model being devised for the institutions of higher education in the State of Pennsylvania essentially differentiates among segments of the population (Newton, 1976). The basic assumption of the model is that an identifiable segment of the general population possesses a set of demographic characteristics that distinguishes it as unique from all others and that associated with each segment is a set of probabilities which describes its distribution among various types of organized educational activities. For this model, the population was divided by sex, race, and age groups. The participation rates of each of these categories were further divided by type of attendance and five different levels of study. For these latter categories they used national figures obtained from NCES and the U.S. Bureau of Census to obtain their participation rates, assuming the State of Pennsylvania would follow national trends. By multiplying the number of each segment of the population by its participation rate, the number of persons projected for enrollment from that category is determined. Then, by adding all the calculated projections together for each population segment for a particular year, they arrive at a forecast. They do, however, have to rely on the accuracy of the population forecast for their projected enrollments to be accurate (Newton, 1976).

A unique and different perspective on enrollments is the approach taken by Aiken Connor (1971) who views enrollments as the product of the demands for educational services by the student population, mediated by the operations of the college. By eliminating those people who cannot attend the college for various reasons, the maximum enrollment the college can expect is obtained. Some of the limiting factors (those factors that eliminate people from attending the college) that would be considered are: all persons enrolled in the public schools, all persons below public school age, persons enrolled in other institutions of higher education, persons who already have a degree equivalent or above that offered by the college, and the admissions policies of the college that might eliminate certain segments of the population. Stimulating factors might be the draft law that exempted certain persons from military service if attending college, the program offerings of the college, and the nature and amount of industry in the area. When all the limiting factors and stimulating factors are applied to the population, a ratio of the actual to the potential enrollees is used as one of the factors to forecast enrollments. Not only is this a unique procedure for forecasting enrollments, but it offers the administration a unique look at the population that it really has some chance of attracting. How well this model would predict enrollments, however, is questionable since Connor (1971) did not test it.

Model Building

The basis of any model is the behavioral equation. Each behavioral equation has three distinguishing characteristics: the nature of the variables, the form of the equation (or equations), and values of the constants (parameters) involved in the equation. Since parameters are usually chosen to force the model to conform to some set of historical data points, the model is usually accurate for the parameter estimation period, but may not be accurate for forecasting the future. The number of behavioral equations determines the intricacy of the model and the type of variables determines the theoretical orientations; i.e., whether the model is based on internal factors or external factors. Two theories exist concerning the appropriate size of a model. Some theoreticians, especially in the field of economics, believe that the more complex and complicated a model, the better chance it has of explaining the behavior of a system since most systems are complex. Some economic models, for example, have as many as 368 equations (Eckstein's DRI model of the entire economic system); other theoreticians believe simplicity is superior since it excludes the exogenous variables which may cloud the effects of important endogenous variables (Ascher, 1978).

Models may be deterministic or stochastic or a combination of the two. A model is deterministic if no chance element operates or where random sequences (chance elements) occur but are scattered randomly above and below zero such that the chance element can be taken as zero. In a stochastic model, the chance element plays a dominant role; i.e., it does not equal zero. The chance element of the last forecast is influenced by the previous chance element and so on down the line. The chance element may also have to be forecast. Deterministic models rarely describe the real world (Gilchrist, 1976).

Accuracy of Forecasting

A forecast can only be considered valid if the model constructed successfully predicts the future. Through an in-depth study of the data, the researcher has chosen the variables and selected the model or models to use. Provided the basic assumptions of the model are valid, a forecast is made. The accuracy of the prediction can obviously be evaluated when data are available for the forecast period. At the time the prediction is made, however, no method exists which will absolutely determine the accuracy of the predictions, but there are certain key statistics which can be developed that are helpful to the researcher in evaluating his prediction. Since all statistics are based on probability theory, even these indicators may be incorrect.

The statistical criteria for successful forecasting are based on the actual errors between the prediction and the observations. When the elementary statistical features of the situation are studied, however, the structure must be assumed stable. These significant statistical features are the mean error, $\overline{e}$, the mean absolute error (MAE), and the mean square error (MSE). The forecast error is defined as the difference between the observation and its forecast; i.e.,

$$e_t = x_t - \hat{x}_t$$

where e_t is the forecast error at some time t, x_t is the observation at some time t, and $\hat{x}_t$ is the forecast for time t. The mean error, $\overline{e}$, is simply the sum of all the forecast errors, e_i, divided by the number of errors considered, n.

$$e = \frac{1}{n}\left(\sum_{i=1}^{n} e_i\right)$$

If the average error is positive, the forecast lags behind the trend; if negative, the forecast leads the trend. The forecast is suspect if the mean error becomes large because this indicates an incorrect analysis of the structure of the data. When the forecast error systematically deviates from zero, the forecasts are said to have a bias. Normally, the mean should be near zero since the actual errors will be both negative and positive and will tend to cancel (Gilchrist, 1976).

The spread of the data can be measured by the mean absolute error or the mean square error. The mean square error places more emphasis on large errors than does the mean absolute error.

$$\text{MAE} = \frac{1}{n}\left(\sum_{i=1}^{n}|e_i|\right) \quad \text{and} \quad \text{MSE} = \frac{1}{n}\left(\sum_{i=1}^{n} e_i\right)^2$$

Both of these methods eliminate the sign from the error terms. For comparability the square root of the mean square error (RMSE) must be used in order to make the units the same. Where a bias is found to exist, the spread should be measured about the average of the error rather than about zero. The sample variance, s^2, would replace the mean square error:

$$s^2 = \frac{1}{n}\left(\sum_{i=1}^{n}\right)\left(e_i - \overline{e}\right)^2 \text{ or } s^2 = \text{MSE} - \overline{e}^2$$

The standard deviation is the square root of the variance. From a statistical point of view, a good forecast can be determined by requiring the average error, the mean square error, or the absolute error to be small.

Summary

Most institutions of higher education are interested in enrollment projections because they are closely related to institutional goals and missions and are, therefore, essential to financial and program planning at every level. In the 1970's the demands on higher education have increased. The general call for "accountability" has caused educators to place much more emphasis on systematic planning. In years prior to 1970 the projection of total enrollments at the national and state level was all that was required because growth rates justified the need for increased resources. In the seventies, however, growth patterns for institutions of higher education became uncertain and new trends did not emerge. The same methods of projecting enrollments in the 1960's did not produce accurate enrollment projections in the 1970's and probably will not produce accurate enrollment projections in the 1980's because enrollment patterns in institutions of higher education are changing. The number of traditional college students are declining while nontraditional students are increasing. As institutions of higher education were expanding their missions and goals to include the new types of students to institutions of higher education, the data base on which enrollment projections were made did not expand and change as rapidly as the growth patterns changed. The assumptions on which enrollment projections were based called for the use of factors for which data were not available. Accuracy in enrollment projections will not increase for institutions of higher education until data is reported for all categories required by the multidimensional educational system which includes both higher education and postsecondary education.

Chapter 3: METHODOLOGY

Purpose

The purpose of this study is to identify the factors relevant for projecting enrollments and to develop such factors into a model that will project enrollments within 1 percent accuracy for a public community college using expo facto data from Thomas Nelson Community College and utilizing the data format developed by the State Council of Higher Education for Virginia. Publicly supported institutions of postsecondary education in Virginia are required to submit 10 year enrollment projections, updated at 1 year intervals. These long range forecasts must include projections of fall headcount enrollment and total FTE enrollment and are used for capital outlay planning (VCCS Policy Manual, 1979). In addition, projections of fall headcount and summer session headcount enrollments, each divided into full-time and part-time categories, and the total FTE enrollment must be submitted to the Virginia State Budget Office (SBO) during the year preceding the even-year session of the Virginia General Assembly for the 2 years of the coming biennium along with actual data for the current biennium. Thus, except for long term planning for capital outlay, the role of enrollment forecasting in the fiscal system adopted by the Commonwealth of Virginia is primarily to provide a basis for financial planning over a short term (up to 3 years). Fall quarter headcount enrollments (sum of full-time and part-time students registered) provide the basis for allocation of classified positions in logistical services and funds required for maintenance and operation of plant facilities. All other budget line items dependent on student enrollment are allocated on the basis of the number of full-time equivalent students enrolled (Appendix M, State Budget Manual). Enrollment projections are also used by the individual community colleges and the VCCS to guide educational, fiscal, and administrative planning (VCCS Policy Manual). Public colleges and universities in Virginia also use these same enrollment projections for internal planning.

The categories of data available for projecting enrollments, therefore, must correspond to those developed by SCHEV and the State Budget Office. The State Legislature in its 1978 Appropriations Act set the limits of accuracy for these projections at ±1 percent.

Hypotheses

If the purposes of this study are to be achieved, it is necessary to test the following hypotheses:

1. Variables can be identified that will produce accurate forecasting models within the limits of accuracy designated by the Virginia State Legislature.
2. Using the variables identified above, forecasting models can be developed for projecting the following categories of enrollment.
    a. Both fall and summer session headcount within ±1 percent
    b. Total FTE within ±1 percent

Selection of Variables

Introduction

As indicated previously, this study involves the development of forecasting models for a comprehensive public community college.

Because every college has its own unique characteristics, a forecasting model must be particularized to an individual college. In this study particular attention is given to projections in enrollment for comprehensive public community colleges within the context of the Virginia Community College System, using Thomas Nelson Community College (TNCC) data. As developed in the literature review, two general types of forecasting models exist, causal and extrapolation models. Development of both types of models requires, as the first step, the selection of independent (predictor) and dependent variables.

Dependent Variables

Dependent variables of any model are determined by the nature of the forecast. In this study, for example, three dependent variables are suggested by the data structure required of the state colleges and universities by SCHEV and SBO. They are:

1. Fall headcount enrollment
2. Summer session headcount enrollment
3. Total FTE

Since each of these categories can be subdivided, several subcategories were also selected for study and are included as dependent variables as follows:

1. First-time and returning student enrollment
2. Part-time and full-time enrollment
3. Male and female enrollment
4. Veteran and nonveteran enrollment
5. Average annual headcount enrollment

Independent Variables

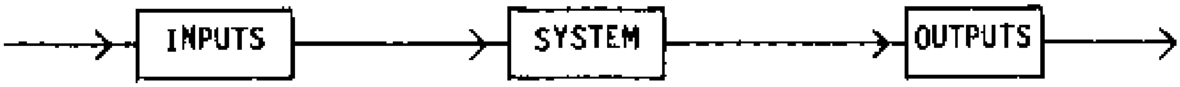
Most forecasting models have specific requirements for the number and kind of independent variables which can be used. Models may be limited by the data available to the researcher. Thus, one of the first tasks in this study was to determine the available data.

Data available for the VCCS falls into two general categories: "census data" and "other data". Only data reported in the VCCS Student Enrollment Booklet (1979) or recorded on magnetic tape in the central systems office of the VCCS in Richmond are considered "census data". Since 1974 the VCCS has had a comprehensive automated system for reporting student enrollment. From 1969 to 1974 the VCCS information system was in the developmental stages and enrollment data were recorded manually at the individual colleges. In 1972 with the first publication of the VCCS Student Enrollment Booklet, enrollment data were standardized. As of that date, enrollment data for all institutions of the VCCS were recorded as of the same census date and the enrollment categories were defined in the same way. Categories are still being added or deleted as the needs change. Age data, for example, did not become "census data" until 1977. "Other data" sources are maintained by the individual community colleges and may or may not be drawn from the same base as "census data". As a result, "census data" are more reliable and accurate than "other data" sources and are comparable for all units of VCCS after 1972.

Data were divided into endogenous and exogenous. Endogenous data are data internal to the system whose behavior is determined by the system itself. Exogenous data are predetermined by some other model or system. Since endogenous data are the most relevant, it usually produces the most accurate forecasts (Gilchrist, 1976). A student flow model was, therefore, constructed to determine the most relevant endogenous data and to understand how students progress through TNCC (see Figure 3.1 on the following page). The possible input and output variables included in the student flow model were then selected as possible independent or predictor variables for the dependent variables using the following criteria: (1) Availability of accurate data, (2) the number of years comparable data are available, and (3) whether the data were from "census" or "other" sources. All input data except for category IV (veteran or nonveteran enrollment) were "census data" and, therefore, available for at least 9 years. Enrollment data for category IV was available for 6 years from the Office of Veterans Affairs at TNCC.

Exogenous independent variables were selected for trial on the basis of one of the following criteria: (1) Variables found successful in previously developed forecasting studies and (2) specific requirements of various models previously discussed in the literature review. Tables 3.1 and 3.2 show demographic and economic data considered for use in this study. The use of the total population in the service area and service area age distribution data were selected as possible dependent variables because NCES (1978) used them with some success in their national enrollment forecasts. Service area high school graduates were selected as a possible independent variable because they are one of the known cohorts for first-time student enrollment. The U.S. enrollment data for all institutions and for 2 year institutions were selected for comparison purposes.

Figure 3.1: Student Flow Model for Thomas Nelson Community College



I. Possible Input Categories for Total Headcount Enrollment Per Quarter

- A. Category I
  - 1. First-time Student Enrollment
    - a. Directly from High School
    - b. Transfer Students
    - c. Other Students
  - 2. Returning Student Enrollment
    - a. From Previous Quarter
    - b. From a Quarter Other than the Previous Quarter
- B. Category II
  - 1. Part-time Student Enrollment
  - 2. Full-time Student Enrollment
- C. Category III
  - 1. Male Enrollment
  - 2. Female Enrollment
- D. Category IV
  - 1. Veteran Enrollment
  - 2. Nonveteran Enrollment

II. The System is Thomas Nelson Community College

III. Outputs

- A. Dropouts Throughout the Quarter
- B. Graduates
- C. Students Completing the Quarter

TABLE 3.1: INDEPENDENT EXOGENOUS VARIABLES - DEMOGRAPHIC DATA

Independent Variables	Years of Availability	Source
Total Population in Service Area	1970 - 78 (Yearly) 1979 - 2000 Estimated in 5-year Intervals	Pop. Projections, Va. Counties and Cities, 1980 - 2000
Service Area Age Distributions	1975 - 82	SCHEV
U.S. Enrollment:		
All Institutions	1968 - 85	NCES (1978 - 79)
2-Year Institutions	1968 - 85	NCES (1978 - 79)
Service Area High School Graduates	1970 - 79 1980 - 82 (Estimated)	Facing Up Local School Board

TABLE 3.2: INDEPENDENT EXOGENOUS VARIABLES - ECONOMIC DATA

Independent Variables	Years of Availability	Source
Local Unemployment	1970 - 79	Va. Employment Commision
Composite Coincident Economic Indicator	1960 - present	Business Condition Digest
Composite Coincident Economic Indicator	1960 - present	Business Condition Digest

In a study by Salley (1978) at Georgia State University, the economic variables listed were successfully used in a time-series decomposition model to project quarterly FTE enrollment. Although Salley suggested exploring local economic data such as the unemployment rate, he did not produce any accurate results using these local economic variables. Table 3.3 gives the composition of the coincident and leading

TABLE 3.3: COMPOSITION OF THE COMPOSITE COINCIDENT AND LEADING ECONOMIC INDICATORS AS PUBLISHED BY NBER IN <u>BUSINESS CONDITIONS DIGEST</u>

I. Composite Coincident Economic Indicator (a composite of 4 coincident economic indicators) - Series No. 920*

1. Employees on nonagricultural payrolls; series no. 41
2. Industrial production (total index); series no. 47
3. Personal income less transfer payments (1972 dollars); series no. 51
4. Manufacturing and trade sales (1972 dollars); series no. 57

II. Composite Leading Economic Indicator (a composite of 12 leading economic indicators) - Series No. 910*

1. Average workweek, production workers, manufacturing (hours); series no. 1
2. Layoff rate, manufacturing (per 100 employees); series no. 3
3. New Orders for consumers goods and materials (1972 dollars); series no. 8
4. Net business formation (index); series no 12
5. Stock prices, 500 common stocks (index); series no. 19
6. Contracts and orders for plant and equipment (1972 dollars); series no. 20
7. New building permits, private housing units (index); series no. 29
8. Vendor performance, percent of companies reporting slower deliveries; series no. 32
9. Net change in inventories on hand and on order (1972 dollars), smoothed with a 4-term moving average (with weights 1, 2,2,-1) series no. 36
10. Change in sensitive prices (percent), smoothed as in 9 above; series no. 92
11. Change in total liquid assets, smoothed as in 9 above; series no. 104
12. Money supply (1972 dollars); series no. 105

* The series numbers refer to the Department of Commerce listings as found in their Business Conditions Digest (BCD) publications.

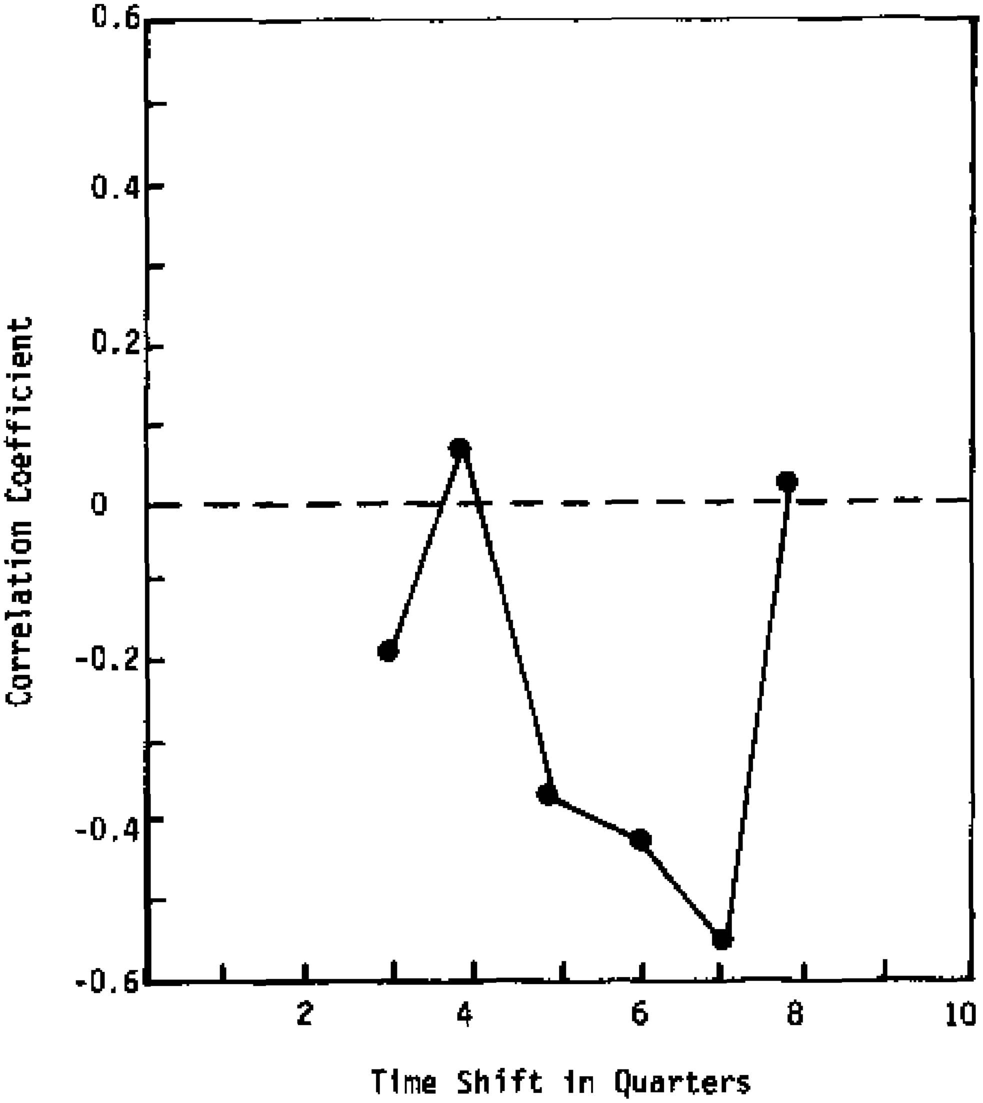
economic indicators which were computed by the National Bureau for Economic Research (NBER) and published in the <u>Business Conditions Digest</u>. The composite coincident and leading economic indicators were the economic variables Salley used to produce his most accurate projections. These indices were also found useful in this study. Attempts were also made to use the local unemployment rate. Which of the economic variables were used and which rejected will be discussed under the section on the Development of Forecasting Models.

Variables were first selected or rejected for inclusion in forecasting models on the basis of simple correlations and the F ratio. Tables IA through IJ, Appendix II show these correlations and the significance level. From these correlations it quickly became evident that most of the endogenous and several of the exogenous variables were strongly related to the dependent variables. Variables not showing strong relationships with the dependent variables were first-time student enrollment, local high school graduates, and the economic indicators.

Since many of the variables with high correlations lacked any time lead with respect to the dependent variables; i.e., they only became available at the time as the dependent variables, they could not be used to predict enrollments using many of the most common forecasting techniques such as linear and multiple regression, and nonlinear growth models. To attempt a forecast using these variables, therefore, required that they be projected. Thus, attempts were made to forecast returning student enrollment because it showed the highest correlations with the dependent variables. These forecasts and the models used for the forecasts are shown in Table E, Appendix III. Since these forecasts had at least a 1-percent error, they were rejected for use at this time in forecasting the dependent variables. A similar result was noted when total population in the service area (S.A.), population, age 18-24 years in S.A., and national two year college enrollments were used to predict the dependent variables. For these three variables, projections had been made by the agencies reporting the data. When these projections were used in forecasting models (in this case, linear regression and nonlinear growth models), the most accurate forecast had an error of 2 percent when compared to actual enrollments. It appears, therefore, that projections based on data which must be projected are unreliable. The errors will be compounded.

When an event occurs, its effects may not be felt immediately, but at some later time. In several cases, it made sense to shift the independent variables forward in time, and investigate the possibility of past events affecting present enrollments. Because part-time students normally attend college for several years, the effects on enrollments for part-time students may be felt for several years. A recession this year may affect enrollments next year or even at a later date. Since part-time student enrollment and economic conditions in previous years might affect present enrollment, these variables were shifted forward in time and correlated with the dependent variables. To determine when the effect betweeen the variables was a maximum, the shift was made for various times. The correlations were then examined to determine when the maximum affect occurred. Figure 3.2 shows the correlations between two variables, one of which was shifted forward in time, three to eight quarters. The maximum correlation coefficient occured when the economic variable (unemployment rate, in this case) was shifted forward seven quarters, indicating its maximum effect on enrollments occurred seven quarters later. The unemployment rate was, therefore, shifted forward seven quarters and later used in a time-series decomposition model.

Figure 3.2: Correlation Coefficient Between Two Variables, With One Variable Shifted in Time



Other variables shifted forward in time and correlated with the dependent variables were first-time student enrollment, and full-time student enrollment. Tables IIA through IIJ, Appendix II show the results of these correlations. The correlations actually increased for first-time student enrollment and the national economic indicators when these independent variables were shifted forward in time. Correlations for other independent variables such as part-time student enrollment and full-time student enrollment decreased, when shifted forward in time, but were still highly correlated with the dependent variables.

Dependent variables were also shifted forward in time and correlated with themselves to determine if prior enrollments were affecting present enrollments. Total headcount enrollment last year, for example, may heavily influence present enrollments. When a variable is shifted in time and correlated with itself, the process is called autocorrelation analysis. Average annual headcount enrollment was such a variable that was highly correlated with itself when shifted forward in time. Table IIA, Appendix II gives the autocorrelations for average headcount enrollment.

Further selection and rejection of the independent variables will be discussed as they are used in specific models in the next section on relevant factors not used in this study. Variables with lower correlations used in various forecasting models because of the specifications of the model will also be discussed in the following two sections.

Relevant Factors Not Used in This Study

Relevant factors used in other studies, and listed in chapter 2, Tables 2.1, 2.2, and 2.3, but rejected in this study are listed in

Table 3.4 to follow along with the reasons for rejection. Table 3.5 lists the variables selected for study and the reasons they were rejected.

Development of the Forecasting Models

Introduction

Forecasting models were selected for investigation for various reasons; availability of data, strength of the simple correlation coefficients, specifications of the models, and trends in the data. Models selected on the strength of the correlations discussed in the previous section were the multiple regression models (including linear and nonlinear relationships) and the student flow model using Markov transition matrices. The cohort-survival model was rejected because of the weakness of the correlation between first-time student enrollment and the local high school graduates. An attempt was also made to use the constant mean model to predict first-time student enrollment because yearly first-time student enrollment had been fairly constant (within ±60 students) since 1972. (See Figure III, Appendix I) When quarterly headcount and FTE enrollment were graphed (See Figure IV, Appendix I), it became evident that a seasonal factor was present. The seasonal variation in headcount and FTE enrollments suggested the use of the time-series decomposition model. These models will be thoroughly discussed in the sections following the section on Statistical Tests for Selecting Forecasting Models. The statistical tests which appeared to have the potential for selecting the most accurate forecasts will be discussed next.

TABLE 3.4: RELEVANT FACTORS USED IN OTHER STUDIES, BUT REJECTED IN THIS STUDY

Factor	Reason for Rejection
1. Age distribution of students institution	Data not available at
2. Type of institution	Not quantifiable
3. Graduate enrollment	Not relevant to institution
4. First professional enrollment	Not relevant to institution
5. Non-degree-credit enrollment	Not considered by state for enrollment projections
6. Pupil to teacher ratios	Not relevant in open admission
7. Military draft	Not relevant at this time
8. Education characteristics of AAP	Local data not available
9. Education characteristics of persons leaving the AAP	Local data not available
10. College attendance rates	Concept not clear
11. Public school enrollment	Not deemed relevant for study
12. Previous college experience of students	Data not available
13. Educational background of parents	Data incomplete
14. Persons from area in other institutions	Data not available
15. Relative educated wage	Data not available
16. Persons in area having degree	Data not available
17. No. of jobs available	Included as unemployment rate
18. Type of jobs available	Service area complex not used this study
19. Jobs in service occupations	Too restrictive for study
20. Relevancy and attractiveness of programs	Data not available
21. Convenience of times and places	Subjective, not quantifiable
22. Admission requirements	Open admissions institutions
23. Tuition charges	Was not applicable to models considered
24. Terms of financial aid	Data not available
25. Release time from work	Data not available
26. Plans of high school students	Data not available
27. Degree of influence of certain persons	Data not available
28. New degrees	Data incomplete
29. Impact of future events	Subjective, not quantifiable
30. Student aspiration	Subjective, not quantifiable
31. Faculty size	Not applicable to models considered
32. Adademic standing	Data not available
33. Type of community	Subjective, not quantifiable
34. Area industry	Subjective, not quantifiable

TABLE 3.5: RELEVANT FACTORS EXPLORED, BUT REJECTED

Factor	Reason for Rejection
1. College going rate	Constant, no correlation
2. Population of service area	Projections produced had large errors
3. Sex of student body	Projections produced had large errors
4. High school graduates	Low correlations with enrollments
5. Birth rate	Did not consider in this study-constant
6. Racial composition	Did not consider in this study-constant
7. Part-time enrollment	Projections produced were not within ± 1%
8. Local unemployment rate	Projections produced were not within ± 1%
9. First-time student enrollment	Projections produced had greater than 5% error
10. Returning student enrollment	Projections produced were not within ± 1%
11. No. of graduates	Projections produced were not within ± 1%
12. No. of dropouts	Projections produced were not within ± 1%
13. Veteran enrollment	No method was found to accurately project

Statistical Tests for Selecting Forecasting Models

The statistical tests used for optimizing the forecasting models were the mean square error (MSE), the average error (e), the multiple correlation coefficient (R), the multiple coefficient of determination (R^2), and the simple correlation coefficient (r), all of which have been examined in chapter 2. The mean square error is a measure of the variation between the actual data and those produced by the model. The average error is an indication of the bias between the forecast and the actual data. The multiple correlation coefficient indicates the quality of the relationship between the dependent variable and the independent variables. The multiple coefficient of determination measures the amount of the variation explained by the model. Thus, if a model produces an accurate forecast the mean square error should be low, the average error should be very close to zero, and the multiple correlation coefficient should be high (as close to 1 as possible). In the context of this study the MSE would need to average 1 percent. For the MSE to average 1 percent, the average difference between the actual enrollment and the forecasted enrollment must be ±1 percent of the actual enrollment. The only truly valid test of a forecast, however, is a comparison with actual data as they become available.

Multiple Regression Models

Probably the most powerful and most widely used methods of statistical forecasting today are multiple correlation and regression techniques. These techniques were investigated because many variables can be quickly explored in many different combinations through their MSE and multiple correlation coefficient. With modern computers, regressor variables (independent variables) can be added and removed from a model and resulting changes examined. Selections of regressor variables for each of the dependent variables was based on the previously discussed simple correlation coefficients and the F ratio.

Multiple correlation and regression techniques were used to determine whether a relationship (linear or nonlinear) existed between independent and dependent variables. The independent variables having the highest correlations with the dependent variables were added and removed from the regression equation on the basis of four criteria: the multiple correlation coefficient, the number of regressor variables indicated, the tolerance level, and the significance level of the F ratio. When the tolerance level (the proportion of the variance not explained by the independent variable already in the regression equation) becomes less than 0.1 percent, no additional independent variables were entered into the equation. The minimum significance level chosen for the F ratio was 0.1 (Nie, 1975). In one instance fifteen regressor variables were selected before the tolerance level and the minimum level for the F ratio were reached. (See Table A, Forecast 2, Appendix II) The number of regressor variables should never exceed the number of observations because the variances in the coefficient become so large that the model is unreliable for forecasting purposes (Gilchrist, 1976). In this study the number of observations for quarterly enrollment projections were limited to 16 or 20 (16 if 1978-79 data were not included) because enrollment patterns changed radically in 1974-75 with the large influx of veterans. When yearly data were used the number of observations were either 10 or 11 (10 if 1978-79 data were not included). In each case, the maximum observation available were used because limiting the number of observations would tend to increase the MSE and reduce the F ratio below an acceptable significance level.

Forecasts using multiple correlation and regression techniques are present in Table A through E in Appendix C. Four of the five most accurate forecasts for fall headcount enrollment were produced using multiple regression analysis. In addition, all of the most accurate forecasts for annual FTE enrollment were produced using this method. The most accurate of these forecasts will be thoroughly discussed in chapter

4. A multiple correlation and regression analysis combined with a time-series decomposition model also produced the most accurate forecasts for fall headcount enrollment.

Nonlinear Growth Models

Leslie and Miller (1974) suggest that enrollment grow exponentially upward in a manner similar to the growth of the economy (see chapter 2). Certain TNCC enrollments such as headcount, FTE, returning student, part-time student, male, and female enrollment appear to grow nonlinearly with time (see Figures II, III, V, VI, VII, Appendix III). Instead of matching the existing data to a linear relationship, nonlinear relationships such as exponential, logarithmic, or power were fitted to the data. Independent variables used were time, part-time enrollment, full-time enrollment, and fall headcount enrollment. Forecasts using the previously listed independent variables are shown in Appendix C with relevant statistics. One forecast using nonlinear regression analysis was within the limits required by the Commonwealth. This forecast will also be discussed in detail in chapter 4.

Markov Transition Matrices Applied to the Student Flow Model

Markov transition matrices applied to student flow models are used by many colleges for projecting enrollments. In a Markov transition matrix, probabilities are calculated based on what a student can be expected to do the following quarter. Two basic assumptions underlie Markov transition matrices: first, the transition between any two states must depend only on the present state and be independent of the the past and second, the system must be stationary; i.e. stable over a period of time. Both assumptions underlying Markov transition matrices are directly violated by TNCC headcount and FTE enrollments. The students at an urban community college do not necessarily return from the previous quarter, but may stop for a quarter or more before returning. Returning student enrollment, therefore, does not depend on the previous quarters' enrollment. The second assumption was violated because the probabilities computed were not stable over time. Table 3.6 on the following page shows the probabilities, the standard deviations, and the standard errors for each probability. The high standard deviation and standard errors indicate the instability of the probabilities.

Although enrollments in general violate the basic assumptions underlying the Markov transition matrix, many colleges have found that good results were obtained using this model. The student flow model previously discussed was used to determine the probabilities needed for the Markov matrix. Probabilities were computed for first-time student enrollment, transfer enrollment, returning student enrollment, the number of graduates, and the number of dropouts.

The probabilities were percentages of the total enrollment in each category averaged for the 4 years, 1974-75 to 1978-79 (see Table 3.6, Part A). When a forecast was attempted using these probabilities, the errors in fall and winter headcount enrollments for 1978-79 are 6.4 percent and 2.8 percent, respectively. A second forecast was attempted using percentages based on the previous years enrollments (see Table 3.6, Part B). The errors for forecasts of summer, fall, and winter enrollments were 1.4 percent, 4.4 percent, and 2.7 percent, respectively. Because forecasts were not within the limits of accuracy required by the

TABLE 3.6: AVERAGE PROBABILITIES FOR THE INPUTS AND OUTPUTS OF THE STUDENT FLOW MODEL USING THE ACADEMIC YEARS 1974-75 TC 1978-79

Input-output	Statistic	Quarters Summer	 Fall	 Winter	 Spring
Part A: Probabilities computed from the total enrollment in a given quarter					
First-time Student Enrollment	p	0.143	0.305	0.132	0.127
	σ	0.023	0.024	0.022	0.022
	Std. Er.	0.011	0.012	0.011	0.011
Transfer Enrollment	p	0.049	0.068	0.037	0.038
	σ	0.024	0.029	0.018	0.020
	Std. Er.	0.012	0.014	0.009	0.010
Returning Student Enrollment	p	0.808	0.628	0.831	0.835
	σ	0.024	0.009	0.004	0.028
	Std. Er.	0.012	0.004	0.002	0.014
Graduates	p	0.027	0.008	0.007	0.064
	σ	0.011	0.030	0.002	0.005
	Std. Er.	0.005	0.016	0.008	0.002
Dropouts	p	0.165	0.256	0.149	0.213
	σ	0.028	0.031	0.013	0.035
	Std. Er.	0.014	0.015	0.006	0.020
Part B: Probabilities computed from the total enrollment the previous quarter					
Returning Student Enrollment	p	0.427	0.726	0.734	0.845
	σ	0.036	0.043	0.028	0.014
	Std. Er.	0.018	0.019	0.019	0.061

Commonwealth, student flow models using Markov transition matrices were rejected as possible forecasting models for TNCC enrollments. The other inputs shown in the student flow model were not used in a Markov transition matrix because no method was discovered to determine what these students would do the following quarter.

The Cohort-Survival Model

Two possible cohorts for first-time student enrollees are service area high school graduates and age groups within the service area population. Accurate forecasts were not possible using either group of cohorts. Between 1975 and 1978 approximately 30 percent of the first-time student enrollees came directly from high school (TNCC Statistical Abstract, 1978-79). The simple correlation coefficient between service area high school graduates and first-time student enrollees was very low at 0.32 and was only significant at the 0.3 level (see Table ID, Appendix II). The only known cohort for the remaining 70 percent of the first-time student enrollees is the age distribution data for the service area population, but age data for TNCC first-time student enrollees were inadequate as previously noted. The cohort-survival model was, therefore, rejected as a possible forecasting model for TNCC enrollments at this time.

Constant Mean Model

A constant mean model was the most accurate forecasting model found for first-time enrollments at TNCC. Since 1972 first-time enrollments at TNCC has been fairly constant, ranging from 813 to 920 students as the yearly average (see Figure III, Appendix I). Using the average over the 7 years as the forecast, an error of 5.8 percent was obtained. All other forecasts produced larger errors (see Table G, Appendix III). Since total enrollment is the sum of first-time and returning student enrollment, these two categories were considered as a method of obtaining a forecast for total enrollment. This method was rejected, however, when it became evident that first-time enrollment could not be projected with any degree of accuracy using this model. Furthermore, the simple correlations between first-time student enrollment and the selected independent variables were much lower than between the other selected dependent variables and the same independent variables (see Table ID, Appendix II). The low correlations (Table III) indicate that other dependent variables should produce better forecasts than first-time student enrollment.

Quarterly first-time enrollment was also examined for forecasting purposes. Although annual first-time enrollment was fairly constant, quarterly first-time enrollment was very erratic (TNCC Statistical Abstract, 1978). No stable pattern was discovered for quarterly first-time enrollment. When first-time fall enrollment dropped, first-time enrollment would increase in the winter and/or spring quarters. When first-time enrollment was high in the fall, it dropped in the winter and/or spring. Thus, over the year the average was fairly constant, but not constant enough to meet the forecasting requirements of the Commonwealth.

Time-Series Decomposition Models

A time-series decomposition model was used by Salley (1979) at Georgia State University for forecast quarterly enrollments. The model assumes that some pattern or combination of patterns are recurring over time. The model is divided into four basic parts: A seasonal factor, a trend factor, a cyclic factor, and an error factor. Patterns may exist when one or more of these factors are not present. The forecast simply combines the parts by addition or multiplication. By using the inverse operation, therefore, the model is decomposed into its parts. The first step is to remove the seasonal factor if one exists. Headcount and FTE enrollment per quarter were seasonalized by computing a centered four quarter moving average. Textbooks such as Wheelwright and Makridalus

(1977) give step by step procedures for seasonalizing data. The ratio of actual data to the moving average was then averaged for each quarter to obtain the seasonal index and then multiplied by the seasonalized data. Seasonalized and actual headcount enrollment are shown in Figure 3.3 on the next page. The plot with large variations is the unseasonalized data. After the seasonal factor is removed, the variations are much smaller. The next factor removed is the trend factor, if one exists. The trend factor is removed by using simple correlation and regression analysis. Essentially the assumption is made that enrollments increase with time. The residual, which is the difference between enrollments predicted by the linear regression equation and the seasonalized enrollments (the dependent variable in the regression analysis), is then correlated with selected cyclic factors. Cyclic factors suggested by Salley (1979) were the economic indicators published by NBER discussed previously, and the local unemployment index. Table 3.7 shows the correlations between the residuals and the indicators shifted in time. The economic indicators were shifted in time because past economic conditions were assumed to affect present enrollments. Present economic conditions should then affect later enrollments. Since the correlations shown in Table 3.7 were very low and the forecast produced (See Table A and D, Appendix III) were not within the accuracy required by the Commonwealth; this part of the model was rejected. What this really implies is that TNCC enrollments do not appear to have a cyclic factor. By combining this model, without the cyclic factor, with a multiple regression model, the most accurate forecast was obtained. A thorough discussion of the most accurate forecasting model will also be found in chapter 4.

Figure 3.3: Headcount Enrollment Per Quarter

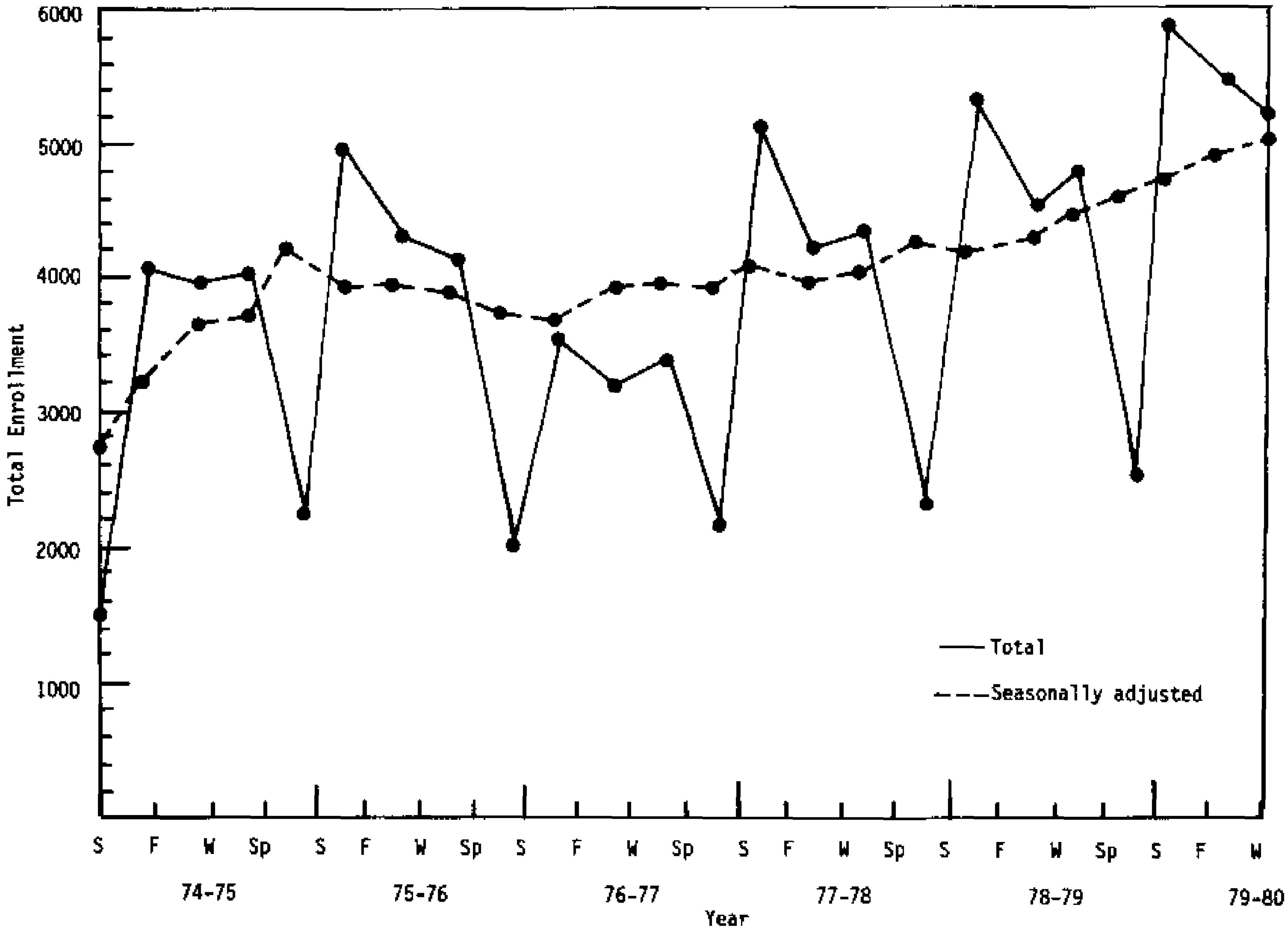


TABLE 3.7: CORRELATION COEFFICIENTS BETWEEN THREE CYCLIC FACTORS AND RESIDUALS IN THE TIME-DECOMPOSITION MODEL

A: Dependent Variables: Total credits seasonally adjusted
Trend Factor: Time

Time	Unemployment Rate		Coincident Indicator		Leading Indicator	
Lead	Difference	Quotient	Difference	Quotient	Difference	Quotient
t+3	-0.1572	-0.1991	-0.0025	-0.0279	0.1672	-0.2021
t+4	0.0783	0.0093	0.1678	0.1487	-0.0834	-0.1150
t+5	-0.3528	-0.3603	0.3872	0.3805	0.1121	0.0842
t+6	-0.4254	-0.4470	0.5924	0.6044	0.4111	0.3991
t+7	-0.5359	-0.5735	0.5432	0.5863	0.5872	0.6068
t+8	-0.0361	-0.0838	0.3163	0.3704	0.4661	0.5126

B: Dependent Variable: Headcount enrollment seasonally adjusted
Trend Factor: Time

Time Lead	Unemployment Rate	Coincident Indicator	Leading Indicator
t+3	0.0172	0.0023	-0.0191
t+4	0.3149	0.0612	-0.0142
t+5	0.0673	0.0123	-0.0012
t+6	0.1681	0.0285	0.0150
t+7	-0.4108	0.0667	-0.0180
t+8	-0.2369	-0.2159	-0.1899

Note: Only the difference was correlated here because the correlation coefficients between the difference and quotients were so very close.

Summary

From this study three types of models were found to produce forecasts which were potentially within the limits of accuracy prescribed: multiple regression, nonlinear growth, and time-series decomposition in conjunction with multiple regression. Further discussion of these models will be found in chapter 4.

The dependent variables were determined by the nature of the study. Using the student flow model, the independent endogenous variables were identified. The independent exogenous variables were identified from past studies or were specified by the model. The independent variables were then selected for inclusion in forecasting models on the basis of the simple corrrelation coefficient and the F ratio. The low correlations between first-time student enrollment and service area high school graduates along with the limited TNCC student age data available eliminated the cohort-survival model from consideration. Because the basic assumptions underlying Markov transition matrices were violated by TNCC enrollments and the error in the forecasts produced by the model were greater than 1 percent, forecasting models based on Markov transition matrices applied to the student flow model were also rejected.

CHAPTER 4: ANALYSIS OF RESULTS

Introduction

In this chapter the optimum forecasts produced for fall headcount enrollment and annual FTE enrollment will be discussed along with the significant statistics and the methodology utilized. A forecast for summer session headcount that is accurate within ±1 percent is also included since forecasts of summer headcount enrollment were very recently added to the requirements of SHEV and SBO. Only the predicted enrollments with ±1 percent were included in this chapter. Forecasts from other models are included in Appendix C.

Fall Headcount Enrollment

A model that utilized a combination time-series decomposition and multiple regression produced the optimum forecast for TNCC's fall headcount enrollment. The original forecast predicted the fall 1978 headcount enrollment within an accuracy of ±1 percent. Two updates for the model have been made: one for fall 1979 and the second for fall 1980. The original forecast and its first update have now been tested against actual enrollments. For fall headcount forecasts, both the original forecast and the first update predicted fall headcount enrollment within the required ±1 percent. The second update is yet to be tested since its projections are for fall 1980 headcount enrollment.

In this forecast equation, the dependent variable was seasonalized headcount enrollment per quarter. (See Time-Series Decomposition Model, Chapter 3, for a description of seasonalizing data.) The independent variables were time expressed in quarters, NBER's national composite coincident economic indicator shifted forward in time 5, 6, 7, and 8 quarters and NBER's national composite leading economic indicator shifted forward 4 and 8 quarters. The degree of shift by the economic variables was determined by the multiple correlation coefficient, the F ratio, and the tolerance level. Shifts from 4 to 10 quarters for each economic indicater were included as possible independent variables. The composition of the economic indicators are given in Table 3.3, Chapter 3. The seasonalized quarterly headcount enrollment along with the independent variables was then used in a multiple regression analysis. Data used in the original forecast included the years 1974-75 through 1977-78; in the first update data were added for 1978-79 and in the second update data for 1979-80 were added. Data for the spring quarter 1980, however, were not yet available for inclusion. The optimum forecasting equations determined in this study are as follows:

Original Forecast:

$$H = 32.01T + 33.77C_1 - 16.32C_2 + 104.38C_3 - 19.55C_4 - 14.30L_1 - 50.82\,L_2 - 884.30.$$

First Update:

$$H = 61.36T + 33.77C_1 - 12.58C_2 + 85.87C_3 + 3.15C_4 - 25.23L_1 - 52.86L_2 - 851.46.$$

Second Update:

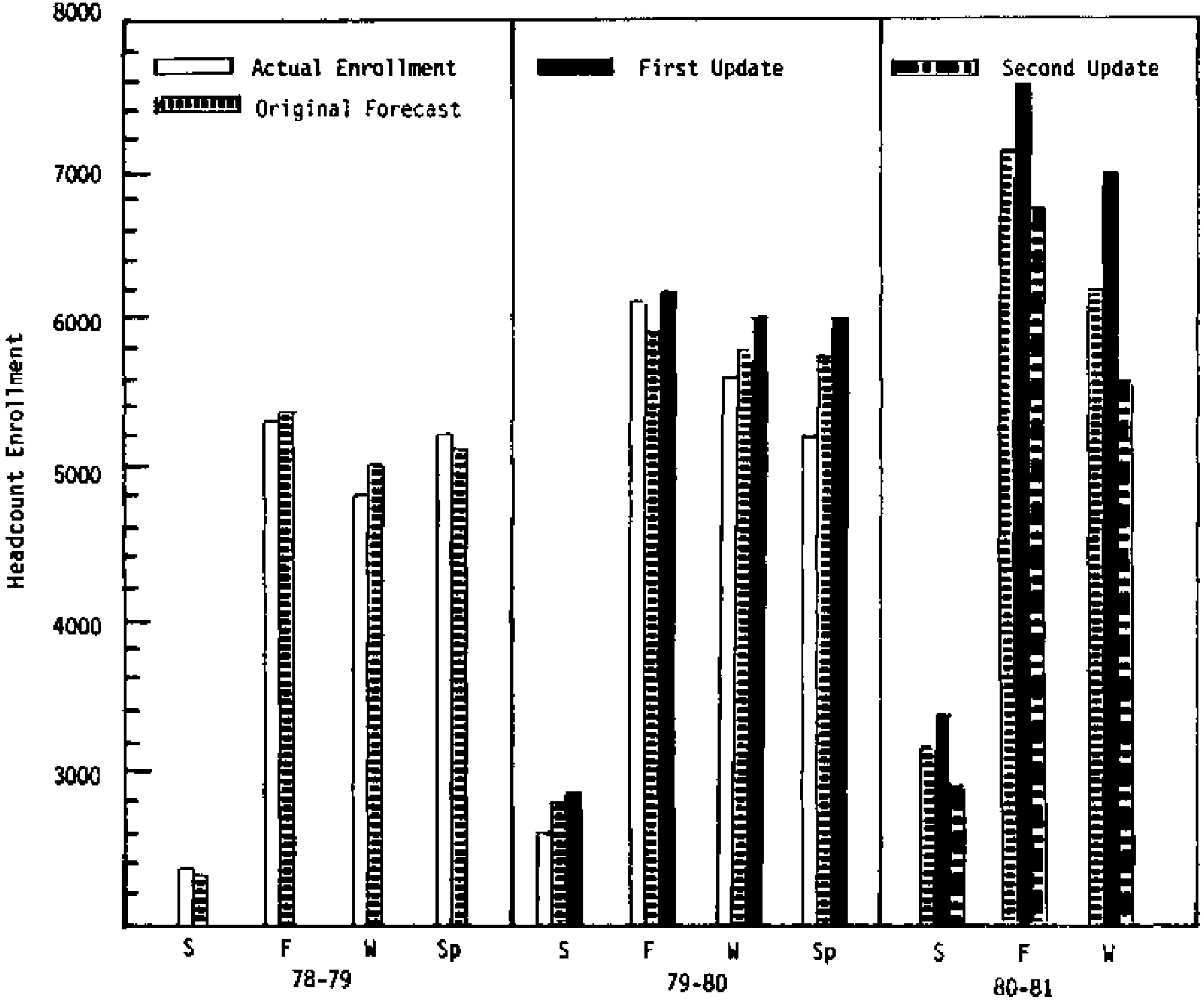
$$H = 67.96T + 38.43C_1 - 39.15C_2 + 51.44C_3 - 17.72C_4 - 28.10L_1 - 5.49L_2 + 3334.75.$$

where, H = seasonalized headcount enrollment; T = time in quarters (T = 1 for summer 1974, T = 2 for fall quarter 1974....$t = n$ for the n^{th} quarter); C_1, C_2, C_3, and C_4 are the national composite coincident economic indicator shifter forward in time 5, 6, 7, and 8 quarters, respectively; and L_1 and L_2 are the national composite leading economic indicators shifted forward by 4 and 8 quarters, respectively. Table 4.1 and Figure 4.1 show the results of the three forecasts.

TABLE 4.1: OPTIMUM FALL HEADCOUNT FORECAST

YEAR	QUARTER	ACTUAL ENROLLMENT	ORIGINAL FORECAST	PERCENT ERROR	FIRST-UPDATE FORECAST	PERCENT ERROR	SECOND-UPDATE FORECAST	PERCENT ERROR
1978-79	Summer	2374	2316	2.4				
	Fall	5312	5316					
	Winter	4728	4895	3.5				
	Spring	4918	4815	2.1				
1979-80	Summer	2579	2719	5.4	2807	8.1	5390	4.5
	Fall	6045	5826	3.6	6069			
	Winter	5517	5745	4.1	5996	8.7		
	Spring	5160	5650	9.4	5955	18.4		
1980-81	Summer		3118		3327		2847	
	Fall		7024		7560		6641	
	Winter		6146		6497		5586	
	Spring							

Figure 4.1: Optimum Fall Headcount Forecast



To place these forecasts in proper perspective, several statistical aspects of the models must be discussd. An error analysis for these models is given in Table 4.2 which includes: the mean square error (MSE), the average error, $\overline{e}$, the multiple correlation coefficient, R, and several other statistical parameters for the original forecast and the two updates. As noted previously, in the ideal case, MSE $\simeq 0$, $\overline{e} \simeq 0$, and R $\simeq 1$. As shown in Table 4.2 for the original forecast, based on 1974-75 through 1977-78 data, since the MSE was 1463 (which is <1 percent) and R = 0.9884, the original forecast is, therefore, considered to be statistically reliable. For the first update, the MSE has degraded significantly to about 8.8 percent and although the model forecasts fall 1979 headcount enrollment within ±1 percent, it does not have the statistical reliability associated with the original model. For the second update, the MSE was also approximately 9.1 percent--it too must be regarded with caution.

It should be noted that these equations can also be used to yield predictions not only for fall headcount enrollment, but also for the winter, spring, and summer quarters. The reader is cautioned, however, that these forecast equations did not meet the predictive criteria for any but the fall quarter. As a special case, a forecast for the summer quarter using the same type model will be discussed in a later section.

In addition to the model just discussed, three other models were developed which would predict fall enrollment within ±1 percent. A linear regression model utilizing only time as the independent variable and unseasonalized fall headcount as the dependent variable was found to predict fall headcount enrollment for 1979. This forecast might prove

TABLE 4.2: ERROR ANALYSIS FOR OPTIMUM FALL HEADCOUNT FORECAST

STATISTICAL PARAMETER	ORIGINAL FORECAST	FIRST-UPDATE FORECAST	SECOND-UPDATE FORECAST
MSE	1463	14940	16294
MSE'*	1622	1697	1793
Std. Dev.	38	122	128
e	4.8	3.9	21.3
R	0.9884	0.9943	0.9930
R^2	0.9769	0.9843	0.9861
% Error in Fall Forecast	0.08	0.4	

*Average MSE for ±1-percent accuracy

to be the best long range predictor and should be further tested. The error analysis for this model is shown in Table 4.3 on the following page. The other two models that predicted fall headcount enrollment were multiple regression models. Both of these models used time as one of the independent variables. In the first case, the other dependent variable was the national composite coincident economic indicator shifted forward 3 years while the second model used the national composite leading economic indicator shifted forward 3 years as its second independent variable. The error analysis for the two models is also shown in Table 4.3. All three of these models had very high MSE's indicating low statistical reliability. In both models it should be noted that the economic indicators utilized were shifted 3 years. In the model for summer session headcount enrollment, to be discussed next, the economic

TABLE 4.3: ERROR ANALYSIS FOR LINEAR REGRESSION MODEL AND MULTIPLE REGRESSION MODELS

	LINEAR REGRESSION MODEL			MULTIPLE REGRESSION MODELS					
INDEPENDENT VARIABLES	TIME			TIME AND COINCIDENT ECONOMIC INDICATOR			TIME AND LEADING ECONOMIC INDICATOR		
FORECAST YEARS	1978	1979	1980	1978	1979	1980	1978	1979	1980
Statistical Parameters									
MSE	66946	71228	65354	61575	71210	65349	66067	62827	57608
MSE'	1223	1367	1553	1223	1367	1553	1223	1367	1553
Std. Dev.	259	267	256	248	267	256	257	251	240
$\bar{e}$	0	0	0	17.3	1.1	0.4	0	0.6	0.4
R	0.981	0.981	0.986	0.985	0.982	0.986	0.981	0.984	0.984
R^2	0.962	0.962	0.972	0.970	0.963	0.972	0.962	0.968	0.975
% Error in Fall Forecast	7.7	0.3		19.1	0.6		5.5	0.4	

indicator utilized was shifted 2.5 years. The question must be asked, what happened 2 to 3 years ago that would affect enrollment? Further discussion on this phenomenon will be found in Chapter 5.

Summer Session Headcount Enrollment

The forecast for summer headcount enrollment was produced while exploring for models to forecast fall headcount enrollment. As previously noted, summer session headcount enrollment is also required by SCHEV. Essentially, the same model used for projecting fall headcount enrollment also produced the forecast for summer headcount enrollment within the ±1-percent accuracy limitation. As for fall headcount enrollment, the dependent variable was seasonalized headcount enrollment. The independent variables for the summer headcount forecast, however, were time and the national coincident economic indicator shifted forward in time by 10 quarters. As with the optimum forecast for fall headcount enrollment, the model was updated twice since the original forecast. The original forecast (summer 1978) and its update (summer 1979) have now been tested against actual enrollments. The original forecast and its first update have predicted summer headcount enrollment within ±1%. The second update is yet to be tested since its projections are for summer 1980. Data used in the original forecast included the years 1974-75 through 1977-78; in the first update data were added for 1978-79 and in the second update data for 1979-80 were added. As in the fall headcount forecast, data for the spring quarter 1980 were not available for inclusion. The forecasting equations determined in the study are as follows:

Original Forecast:

$$H = 49.13T + 26.89C + 123.01.$$

First Update:

$$H = 47.91T + 17.08C + 103.38.$$

Second Update:

$$H = 51.48T + 24.82C + 339.88.$$

where H is the seasonalized headcount enrollment, T = time in quarters, and C is the national composite coincident economic indicator shifted forward in time by 10 quarters. Table 4.4 and Figure 4.2 show the results of the three forecasts.

The error analysis for this model is given in Table 4.5. Although the model accurately predicts summer headcount enrollment accurately for 2 consecutive years, the MSE indicates that the forecasts should be regarded with some caution. For the original forecast, the first update and the second update, the MSE's showed an accuracy of 21 percent, 17 percent, and 16 percent, respectively. It is interesting to note that the forecasts for the other quarters are not within the prescribed accuracy. This was also true for the optimum fall headcount. Perhaps it suggests that different combinations of variables will be needed to accurately (within ±1 percent) forecast the other quarters, but since SCHEV and SBO only require fall and summer session headcount forecasts, accurate forecasts were not investigated in detail for the winter and spring quarters at this time.

FTE Enrollment

The optimum forecast for total FTE enrollment was produced by a multiple regression model. The independent variables were full-time

Figure 4.2: Summer Headcount Forecast

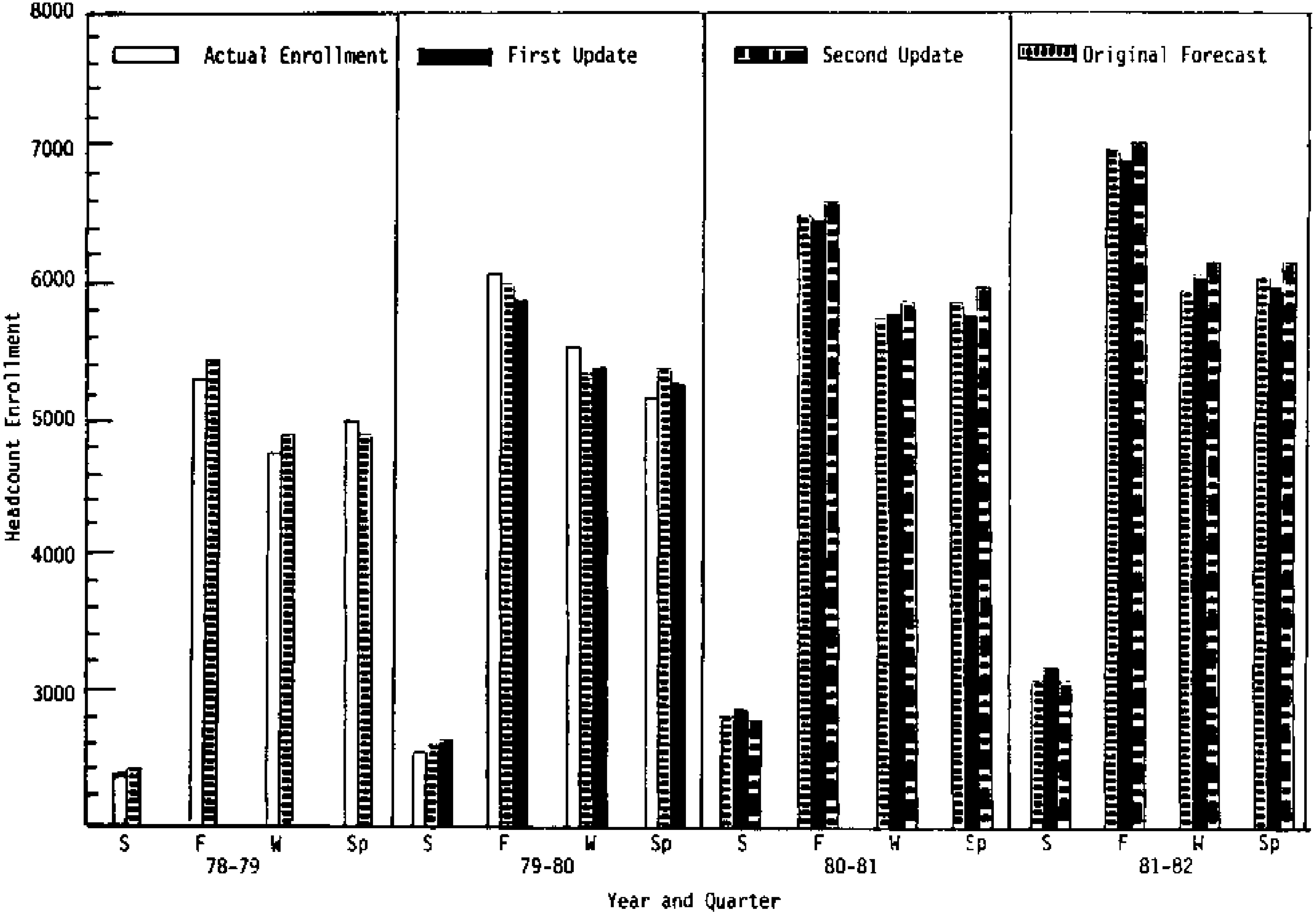


TABLE 4.4: SUMMER HEADCOUNT FORECAST

YEAR	QUARTER	ACTUAL ENROLLMENT	ORIGINAL FORECAST	PERCENT ERROR	FIRST-UPDATE FORECAST	PERCENT ERROR	SECOND-UPDATE FORECAST	PERCENT ERROR
1978-79	Summer	2374	2386	0.5				
	Fall	5312	5420	2.0				
	Winter	4728	4818	1.9				
	Spring	4918	4869	1.0				
1979-80	Summer	2579	2596	0.7	2602	0.9		
	Fall	6045	5909	2.2	5863	3.0		
	Winter	5517	5266	4.5	5290	4.1		
	Spring	5160	5347	3.6	5278	2.3		
1980-81	Summer		2819		2815		2808	
	Fall		6470		6420		6561	
	Winter		5747		5771		5851	
	Spring		5842		5765		5954	
1981-82	Summer		3084		3090		3082	
	Fall		6883		6826		6992	
	Winter		6074		6096		6194	
	Spring		6108		5988		6193	

TABLE 4.5: ERROR ANALYSIS FOR SUMMER HEADCOUNT ENROLLMENT

STATISTICAL PARAMETER	ORIGINAL FORECAST	FIRST-UPDATE FORECAST	SECOND-UPDATE FORECAST
MSE	34670	29469	27029
MSE^F	1622	1697	1793
Std. Dev.	184	170	164
e	28.5	23.7	0.2
R	0.985	0.97	0.99
R^2	0.97	0.94	0.98
% Error in Summer Forecast	0.5	0.7	

enrollment and the national composite coincident economic indicator shifted forward 3 years. The original forecast and the first update have both been tested against actual FTE enrollments. In both instances, the errors between the forecast and actual enrollments for 1978-79 and 1979-80 were 0.03 percent and 0.05, respectively. The second update is yet to be tested since its projections are for 1980-81 FTE enrollments. The forecast equations are as follows:

Original Forecast (1978-79):

$$FTE = 1.01F + 37.93C - 2885.95$$

First Update (1979-80):

$$FTE = 1.01F + 37.93C - 2888.34$$

Second Update (1980-81):

$$FTE = 1.01F + 37.89C - 2874.59$$

where FTE is the total FTE enrollment for 1 year, F is full-time student enrollment shifted forward 3 years, and C is the national composite coincident economic indicator shifted forward 3 years. Table 4.6 shows the forecasts and actual enrollments, and the percentage of error.

When the statistical tests are applied to the optimum FTE forecast, the basis for caution emerges. As may be seen in Table 4.7, the MSE's are extremely high (over 100 percent) indicating the possibility of a large variability existing between the forecasts and actual FTE enrollment. The multiple correlation coefficients are lower than might be desirable, and the multiple coefficients of determination indicate that at least 20 percent of the variation is unexplained by the variables utilized in the forecast. The average errors indicate that there is no bias in the forecasting model. In conclusion, this model produced the optimum forecast on the basis of the errors produced between the forecast and actual enrollments. When the statistical tests are applied, however, the model must be viewed with caution.

Only one other model (Forecast 3, Table C, Appendix III), a nonlinear growth model produced forecasts with less than 1-percent error. In this case, the independent variable was average yearly headcount enrollment shifted forward 1 year. In the original forecast, the error between the forecast and actual enrollments was 0.8 percent. In the first update, the error had increased to 3.3 percent.

TABLE 4.6: OPTIMUM TOTAL FTE FORECAST

Year	Actual Enrollment	Original Forecast	Percent Error	First Update Forecast	Percent Error	Second Update Forecast
1978-79	3229	3230	0.03			
1979-80	3470	3437	0.04	3437	0.4	
1980-81						3720

Figure 4.3: Graphical Representation of FTE Forecast and Actual Enrollment

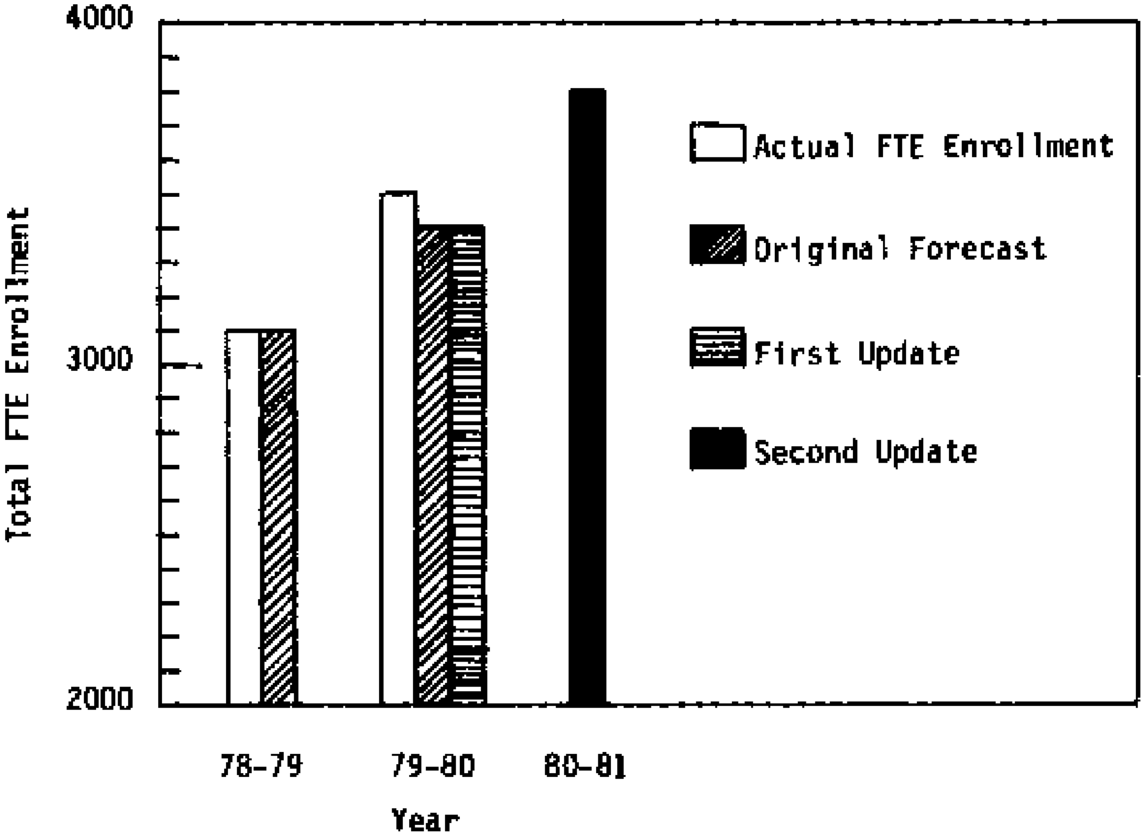


TABLE 4.7: ERROR ANALYSIS FOR OPTIMUM TOTAL FTE ENROLLMENT FORECASTS

STATISTICAL PARAMETER	ORIGINAL FORECAST	FIRST-UPDATE FORECAST	SECOND-UPDATE FORECAST
MSE	103,598	90,589	80,613
MSE'	573	619	775
Std. Dev.	322	301	284
$\overline{e}$	-0.1	0	0.8
R	0.8847	0.8723	0.8936
R^2	0.7827	0.7609	0.7986
% Error in Fall Forecast	0.03	0.4	

The MSE for this forecast was lower at 35 percent than for the optimum FTE forecast. The multiple correlation coefficient for this forecast was 0.98, again much higher than for the optimum forecast. The statistical tests for this model indicate a much more reliable forecast than the statistical tests for the optimum forecast. Forecasts were considered optimum when the errors between the forecast and actual enrollment were a minimum.

Discussion

The fall and summer headcount forecast models can be classified as both an extrapolation and causal model. The time-series portion of these models is considered an extrapolation because it assumes some pattern or combination of patterns are recurring over time. Two patterns are present in the optimum forecasts for fall headcount enrollment and summer session headcount enrollment, the seasonal pattern and the trend pattern. The trend was time-series. The seasonal pattern was evident from examining the repetitive nature of headcount enrollment over the four quarters of successive sessions (see Figure 3.3, Chapter 3). Once the seasonal factor is removed, the trend factor becomes evident, especially from 1976 to present. Time and seasonal factors by themselves, however, did not produce sufficiently accurate forecasts (see Table A, Forecast 6, Appendix III). Since no periodic swings appeared when the seasonal factor was removed, a cyclic factor did not appear to exist. Furthermore, when the residuals from the time-series portion of the model were correlated with economic factors, the correlation coefficients were very low. This also indicated the absence of the cyclic factor (see Table 3.7, Chapter 3). When the time-series portion of the model is combined with multiple regression, a forecast was produced within ±1 percent. The multiple regression portion of the model is causal. The cause-effect relationship between quarterly headcount enrollment and economic factors is somewhat obscure, however, because the economic indicators which produced the most accurate forecasts were shifted forward in time from 4 to 10 quarters. The simple correlations between these economic indicators and enrollment were much lower than for other independent variables. When these economic indicators were shifted in time, however, the simple correlations did improve. An explanation for the shift in time might involve the different age groups enrolled at TNCC. Since the students range in age from 17 to 70, various age groups may be affected differently by economic conditions. When economic conditions worsen, for example, it may be that the younger people go back to school, whereas the older students "stop out". This cause-effect relationship could not be tested at this time because age data have not been available.

The optimum forecasts of total FTE enrollment were produced by a multiple regression model. The independent variables were the national composite coincident economic indicator shifted forward 3 years and full-time student enrollment was shifted forward 3 years. As with the models for fall and summer session headcount enrollment, the cause-effect relationship is obscure because the variables are shifted forward in time. No explanation could be found for this particular time shift. It is noteworthy, however, that three of the other forecasts that met the ±1 percent accuracy criteria also had variables shifted forward 2.5 or 3 years.

Although many of the independent variables selected for study did not enter into the actual forecasting equations, apparent cause-effect relationships exist between several independent and dependent variables and are indicated by the high simple correlations produced. Veteran enrollment, for example, strongly impacted enrollments between 1974 and 1978. It correlated highly with average annual headcount enrollment, fall headcount enrollment, total FTE enrollment, and returning student enrollment. When 1978-79 data were included, the correlations were much lower indicating that veteran enrollment is not having the same effect on enrollments now that it did prior to 1978.

Other independent variables having high correlations with the dependent variables were returning student enrollment, part-time student enrollment, full-time student enrollment, male enrollment, female enrollment, population in the service area (age 18-24 years), and total population in the service area. Obviously these independent variables affect enrollments. No accurate method was developed that allowed utilization of these variables, however.

The realiability of projections is determined by the stability of the factors used in the forecast. Statistically based forecasting models can only project within the limits of the stability. Those responsible for forecasting enrollments, therefore, must continually examine the data for changes that will destroy this stability. Examples of changes at TNCC may in the future--or have in the past--destroyed this stability are:

1. The large influx of veterans in 1974 and 1975
2. Opening new campuses in 1978 and 1979
3. Discontinuing scheduled classes at an earlier date because of insufficient enrollment, a matter of administrative policy
4. Discontinuance of federally financed programs within the institution

Factors such as those listed above make both a positive and negative impact on enrollments. Identifying these factors and the direction and extent of the impact by the decision makers may well determine the accuracy of any final forecast.

Test of Hypotheses

Hypothesis.- Variables can be identified that will produce accurate forecasting models within the limits of accuracy designated by the Virginia State Legislature.

Independent (predictor) variables utilized in the enrollment forecasting models were identified through the literature review and the student flow model. The optimum forecasts (which will be discussed under hypothesis 2) were within the limits of accuracy. The variables utilized in these models were selected on the following basis:

1. Simple correlations and the F ratio
2. Repeated examinations of the correlation between the fitted regression equation and the dependent variable; i.e., by examining the multiple correlation coefficient
3. By examining the changes produced when variables are added and removed from the model; i.e., by examining the MSE, $\overline{e}$, the F ratio for R, and the tolerance level (the percentage of unexplained variance)

The following independent variables were utilized in the optimum forecast models of the three major dependent variables:

For the optimum fall headcount forecast:

1. Seasonal factor
2. Time trends
3. The national composite coincident economic indicator shifted forward 5, 6, 7, and 8 quarters
4. The national composite heading economic indicator shifted forward 4 and 8 quarters

For the summer session headcount forecast:

1. The seasonal factor
2. Time trend
3. The national composite coincident economic indicator shifted forward 10 quarters

For the optimum total FTE forecast:

1. Full-time head count enrollment shifted forward 3 years
2. The national composite coincident economic indicator shifted forward 3 years

For fall headcount enrollment, forecasts were accurately produced for 1 year and for summer session headcount and total FTE enrollment, forecasts were accurately produced for 2 years within 1 percent. On the basis of very short range forecasts (≤ 2 years), hypothesis 1 is accepted because independent variables were identified and utilized in forecasts within the prescribed accuracy. Although independent variables were identified and utilized in longer range forecasts (>2 years), the forecasts were not within the required accuracy. Thus, on the basis of forecasts for >2 years, hypothesis 1 is rejected.

Hypothesis 2.- Using the variables identified in hypothesis 1, forecasting models can be developed for projecting the following categories of enrollment:

1. Fall and summer session headcount within ±1 percent
2. Total FTE within ±1 percent

The basis for acceptance or rejection of hypothesis 2 was made in context with the fiscal system adopted by the Commonwealth of Virginia for the distribution of public funds to the state colleges and universities. In the Commonwealth of Virginia, enrollment projections for the state colleges and universities are required before funds are appropriated. Projections for fall headcount, summer session headcount, and total FTE enrollments must be submitted during the year preceding the even year session of the Virginia General Assembly for the 2 years of the coming biennium. Updates of these projections may be made or even required annually. Thus, forecasts for the 1982-84 biennium are based on 1979-80 enrollment data. Since projections must be made using data 3 years prior to the forecasts, forecasting models must project enrollments for 3 years within ±1 percent.

Although the Virginia state colleges and universities are permitted to update the 3-year forecasts, the original distribution of funds for the biennium is based on those forecasts. Basically, the updates are used to adjust the distribution of funds when actual enrollments are more than 1 percent over or under projections for that session.

Forecasting models for fall and summer session headcount and total FTE enrollment were developed that projected enrollments for 1 or 2 years within ±1 percent accuracy. The optimum fall headcount forecast, a combination time-series decomposition and multiple regression model, has been updated twice. The original forecast and the first update have been tested against actual enrollments and found to accurately project fall headcount enrollment for 1 year within ±1 percent. The second update has not been tested against actual enrollments since its projections are for fall 1980. The statistical tests (MSE and $\overline{e}$) for the original forecast were excellent but deteriorated somewhat for the two updates although the projections from the updates still were within the limits of accuracy required.

The summer session headcount forecast, also a combination time-series decomposition and multiple regression model, has been updated twice. The original forecast and the first update have been tested against actual enrollments and found to accurately project summer headcount enrollment for 2 years to within ±1 percent. The second update is yet to be tested. Some caution must be exercised with the summer session headcount forecast because the statistical tests (MSE and $\overline{e}$) indicated possible variability.

The optimum forecast for total FTE enrollment, a multiple regression model, has also been updated twice. The original forecast and first update have been tested against actual enrollments and found to accurately project total FTE enrollment for 2 years to within ±1 percent. The second update cannot be tested until spring 1981. Caution must also be exercised in the utilization of this model because the statistical tests indicated possible variability between the forecasts and actual enrollments.

The forecasting model produced for fall headcount enrollment can be utilized for the yearly updates which may be required by the State Council of Higher Education of Virginia or needed for adjustment in projecting enrollments. The summer session headcount and FTE enrollment forecasting models can be utilized for the yearly updates and projections for the first year of the biennium. On the basis of the yearly updates, hypothesis 2 is accepted for fall and summer session headcount enrollment projections. On the basis of projections for the first year of the biennium, hypothesis 2 is accepted for summer session headcount and total FTE enrollment projections.

Hypothesis 2 is rejected for forecasts beyond 2 years for fall and summer session headcount and total FTE enrollment projections.

Forecasting models which projected for more than 2 years were not within the prescribed accuracy (see Appendix C).

Summary

Forecasting models for fall headcount enrollment projections for 1 year were produced to within ±1 percent; forecasting models for summer session headcount and total FTE enrollment were produced for 2 years to within ±1 percent. Thus, both hypotheses 1 and 2 were accepted for short range forecasts ($\leq$2 years). Both of the hypotheses were rejected for longer range forecasts (>2 years).

Forecasts for yearly projections up to 10 years are also required by SCHEV. Since these forecasts may be updated periodically, the accuracy is not as critical as for the yearly updates and the forecasts for the biennium. Forecasting models which project fall and summer session headcount and total FTE enrollments for 10 years were produced for this study but none were within the prescribed limits of accuracy.

CHAPTER 5: SUMMARY AND IMPLICATIONS FOR FUTURE RESEARCH

Summary

The State Legislature of Virginia gave to SCHEV the responsibility for distributing the public funds for the institutions of higher education of Virginia fairly and equitably. To provide a basis for such a distribution, the institutions of higher education were required to project their enrollments to within ±1 percent limits of accuracy.

The purpose of this study was to identify the factors relevant for projecting enrollments and to develop such factors into a model that will project enrollments to within 1 percent accuracy for a public community college using expo facto data from TNCC and utilizing the data format of the State Council of Higher Education for Virginia.

To achieve these purposes, the following hypotheses were tested:

1. Variables can be identified that will produce accurate forecasting models within the limits of accuracy designated by the Virginia State Legislature
2. Using the variables identified above, forecasting models can be developed for projecting the following categories of enrollment:
    a. Fall headcount and summer session headcount to within ±1 percent
    b. Total FTE to within ±1 percent

Two general types of statistical forecasting models exist--causal and extrapolation. Development of both types of models require, as the first step, the selection of independent (predictor) and dependent variables. The dependent variables were determined by the nature of the forecast, in this case, by the structure required of the state colleges and universities by SCHEV and the Virginia SBO. They were:

1. Fall headcount enrollment
2. Summer session headcount enrollment
3. Total FTE enrollment

Independent variables utilized in the enrollment forecasting models were identified through the literature review and from the student flow model for the institution. Optimum forecasts for the previously listed dependent variables were produced by time-series decomposition and multiple regression models. The independent variables utilized in these models were selected on the following basis:

1. Simple correlations and the F ratio
2. Repeated examinations of the correlation between the fitted regression equation and the dependent variable; i.e., by examining the multiple correlation coefficient
3. By examining the changes produced when variables are added and removed from the model; i.e., by examining the MSE, $\overline{e}$, the F ratio for R, and the tolerance level (the percentage of unexplained variance)

The following independent variables were utilized in the optimum forecasting models:

For the optimum fall headcount forecast:

1. The seasonal factor
2. Time trend
3. The national composite coincident economic indicator shifted forward 5, 6, 7, and 8 quarters

4. The national composite leading economic indicator shifted forward 4 and 8 quarters

For the summer session headcount forecast:

1. The seasonal factor
2. Time trend
3. The national composite coincident economic indicator shifted forward 10 quarters

For the optimum total FTE forecast:

1. Full-time enrollment shifted forward 3 years
2. The national composite coincident economic indicator shifted forward 3 years

Independent variables which were rejected on the basis of the simple correlation coefficients and the F ratio were service area high school graduates and first-time student enrollment. Other independent variables such as returning student enrollment and veteran enrollment were rejected because no method was found to project them within the prescribed accuracy. Since age range data for TNCC students only became available in 1977, the age data available were considered insufficient to serve as a basis for a reliable forecast at this time.

Forecasting models (within ±1 percent accuracy) were produced for 1 year for fall headcount enrollment and for 2 years for summer session headcount and total FTE enrollment. For fall and summer session headcount enrollment, a model combining time-series and multiple regression produced the optimum forecasts. For total FTE enrollment, the optimum forecast model was a multiple regression model. Each model has been updated twice; for the first update, the projections produced for each model were within the prescribed accuracy. The second update has not been tested because actual enrollment data are not yet available. Discussion of the hypotheses will be given in the next section.

The basis for acceptance or rejection of the hypotheses was made in context with the fiscal system of the Commonwealth of Virginia for the distribution of public funds to the state colleges and universities. The fiscal system was established primarily to provide a basis for financial planning over 3 years. The state colleges and universities must submit fall and summer session headcount and total FTE enrollment projections to the Virginia SBO during the year preceding the even year session of the Virginia General Assembly for the 2 years of the coming biennium. In addition, the state colleges and universities must submit 10-year enrollment projections for the same categories of enrollment to SCHEV. All projections may be updated periodically.

Forecasting models were produced for 1 year for fall headcount enrollment and for 2 years for summer session headcount and total FTE enrollment to within ±1 percent. The forecasting model produced for the fall headcount enrollment can be utilized for the yearly updates which may be required by SCHEV or needed for adjusting the enrollment projections for the upcoming biennium. The summer session headcount and FTE enrollment forecasting models can be utilized for the yearly updates and projections for the first year of the biennium. On the basis of the yearly updates, both hypotheses were accepted for projections of fall and summer session headcount enrollment. On the basis of projections for the first year of the biennium, both hypotheses are accepted for summer session headcount and total FTE enrollment projections. Both hypotheses are rejected on the basis of forecasts beyond 2 years for fall and summer session headcount and total FTE enrollment projections. Forecasts for yearly projections up to 10 years which are required by SCHEV were produced, but not within the prescribed accuracy. Since the 10-year forecasts may be updated periodically, the accuracy is not as critical as for the yearly updates and the forecasts for the biennium.

In conclusion, very short range forecasts (≤2 years) can be produced within the prescribed accuracy for fall and summer session headcount and total FTE enrollments. Longer range forecasts (>2 years) were not produced within the prescribed accuracy.

Conclusions

Based on a review of past enrollment projection studies at the national, state, and institutional levels, certain basic conclusions were reached which should be applicable to postsecondary enrollment projections in general:

1. The large variations between various enrollment projections at the national level have occurred because the forecasts have been based on different assumptions. Since forecasters are essentially examining the same data base (the U.S. Bureau of Census), the assumptions on which the forecasts are based are primarily responsible for differences among the various forecasts.
2. Enrollment projections have been made for institutions of higher education with diverse goals and missions. Separating institutions of higher education into categories of institutions with similar goals and missions may facilitate the isolation of relevant factors. Factors relevant for one type institution of

higher education are not always relevant to all institutions of higher education.

3. Present data bases are inadequate to produce accurate enrollment projections for the following reasons:
   a. Patterns of student enrollment vary among geographical areas and among types of institutions.
   b. There are inadequate data related to patterns of enrollment for various age groups, especially beyond 24 years of age
   c. Data related to the minority segment of the student cohort are inadequate.
4. Most enrollment projections are not sufficiently reliable for planning purposes. Because key statistics, such as the mean square error, the average error, and multiple coefficient of determination, have not been reported for individual studies, the overall potential accuracy of projections cannot be evaluated statistically prior to the time that "hard data" become available.

Since there is a need for accurate enrollment projections at all levels of postsecondary education, an in-depth study for one component of the higher education system (a multidimensional urban community college) was conducted to determine what problems existed in producing enrollment forecasting models within the accuracy limitations prescribed by one state (Virginia). Based on this study, the following set of general conclusions were reached:

1. It appears that accurate projections of enrollment for community colleges will require use of data in projections not

currently classified as census data by the various state coordinating boards.

2. When the impact of the seasonal variations in enrollment among quarters is statistically removed, the identification of relevant factors becomes less difficult.
3. Traditional projection models, utilizing cohort survival and Markov transition matrices, are not applicable in community colleges with open admissions and a nontraditional student population.
4. Accurate statistical forecasting models, using techniques such as multiple regression or time-series decomposition, can be developed which will be accurate for periods of nearly 2 years, but accuracy becomes questionable for longer periods of time.
5. Since Virginia institutions are required to make enrollment projections which are considerably longer than 2 years for biennial budgets, it appears that the current limits of accuracy of ±1 percent for such projections are unrealistic. However, annual updates of biennial projections should be possible within these limits.
6. Results of this study support the concept that application of statistical techniques for verification of the accuracy of prediction models is valuable in evaluating the potential accuracy of new models.
7. In institutions with relatively unstable enrollments, such as community colleges, the factors impacting the total enrollment are constantly in flux; thus, the models for projecting

enrollments in such institutions must be constantly evaluated and/or reformulated to identify and incorporate changing trends as they occur.

8. This study suggests that male enrollment has reached its zenith (see Figure VII, Appendix A). Female enrollment is continuing to increase but at a slightly lower rate. This may indicate that female enrollment is also approaching its zenith. Female enrollment passed male enrollment in 1978. Glenny (1980) also observed the same pattern in national enrollment trends.

In Chapter 2 the relevant factors identified in previous research involving enrollment projections were grouped into three major categories: demographic factors, economic factors, and institutional factors. Each of these categories is briefly discussed in terms of the findings of this study:

1. The traditional demographic factors involved in the research of Cartter (1973), NCES (1978), and Carnegie (1978) have not proved to provide an adequate base for accurate projections of enrollment. More recent research, such as the work of Glenny (1980), indicates a need for redefinition of these factors. The results of this study seem to support this concept in such areas as student age distribution, enrollment of minorities, and the changing proportion of men and women in the student body.
2. The research of Dresch (1975), Froomkin (1974), and Leslie and Miller (1974) utilized economic factors in enrollment largely based upon employment and employment opportunities. Salley

(1978) added the use of the NBER economic indicators as predictors of enrollment change. The models developed for this study utilize these economic indicators, thereby, supporting Salley's research and broadening its applicability to include community colleges with open admissions.

3. Two findings related to the use of institutional factors were reflected in the findings of this study:
    a. The traditional student flow models of Wasik (1971), Lightfield (1975), and others are not appropriate for institutions with open admissions and a large proportion of part-time students.
    b. This study provided strong support for the research of Norris (1976) and others that each institution must constantly study and reevaluate the relative impact of all relevant factors on the enrollment of the institution.

Implications for Future Research

New enrollment projection studies at all levels of higher education are an essential part of institutional planning. As the importance of effective institutional planning is recognized, the necessity of continuous processes of updating, revising, and exploring new techniques of projecting enrollments also become evident. Forecasting models which give reliable results now may not be effective for institutional planning in the future because enrollment patterns are changing. New and different types of students are attending college, especially community colleges and multidimensional universities with open admissions policies. Factors now relevant to enrollments are likely to change in the future. For these reasons, new studies are needed. Specifically, more studies are indicated at this comprehensive urban community college to continually update the existing models and to increase the reliability of the models.

Enrollment projection studies for institutions of higher education should be categorized to determine the factors relevant to specific types of institutions. It is suggested that categories considered comparable include institutions with similar goals and missions.

Because key statistics, such as the mean square error, the average error, and multiple coefficient of determination have not been reported in previous enrollment projections, the overall accuracy of these studies cannot be evaluated for accuracy and reliability by higher education planners. It is suggested that in the future these key statistics be derived and reported for all enrollment projection studies.

Not only have enrollment forecasters given little attention to the statistical accuracy and reliability of their projections, but higher educational planners have not developed a clear picture of their accuracy requirements for enrollment projections. A possible area of future research is clearly a study of the accuracy of enrollment predictions required by higher educational planners at all levels.

New studies should also include an investigation of age data. Possible relationships might exist between age groups and economic factors. The younger students, for example, might go to college when poor economic conditions exist, such as a poor job market; older students might drop out under these same conditions because they could not afford to attend college. In a poor job market, both work situations have been observed; i.e., both increasing and decreasing enrollments. Age data first became available as census data for Virginia institutions in 1977-78. At the time of this study, therefore, insufficient age data were available to investigate existing relationships between it and the independent variables selected for study.

A number of other specific studies to improve the accuracy of enrollment predictions for open multidimensional urban community colleges have been suggested by this work. Among these are: improved forecasts for graduates of high schools in the service area (cohort-survival techniques have traditionally been very successful at the high school level), in-depth examination of attendance patterns and personal characteristics of "stop-outs" (students who leave college for a term or more but return) as opposed to "drop-outs" (students who never return), and improved understanding of the motivation and attendance patterns of nontraditional students who enroll for one or more terms primarily to improve their academic or vocational skills. Veteran enrollment constitutes still another relevant factor which has significantly influenced higher educational enrollment in three instances over the last 35 years--and may still--as a large source of error in projecting enrollments. Finally, this work strongly suggests that researchers closely examine economic factors related to higher education enrollment to establish specific cause and effect relationships in order to give credibility to their use in forecasting.

In conclusion, accurately projecting higher educational enrollments is likely to become a difficult process in the future. Economic factors could assume more importance than they have in the past and trends in enrollment patterns may not remain stable for long periods of time. Nontraditional students, minorities, and women with nontraditional motivational patterns are likely to become more important elements of an institution's enrollment, particularly in multidimensional institutions with open admissions policies such as community colleges. These trends will place increased emphasis on more precise planning at all levels in higher education and, thereby, place increased emphasis on accurate prediction of enrollments.

APPENDIX I

GRAPHS PERTAINING TO ENROLLMENTS

Figure I: Total Enrollment in Institutions of Higher Education

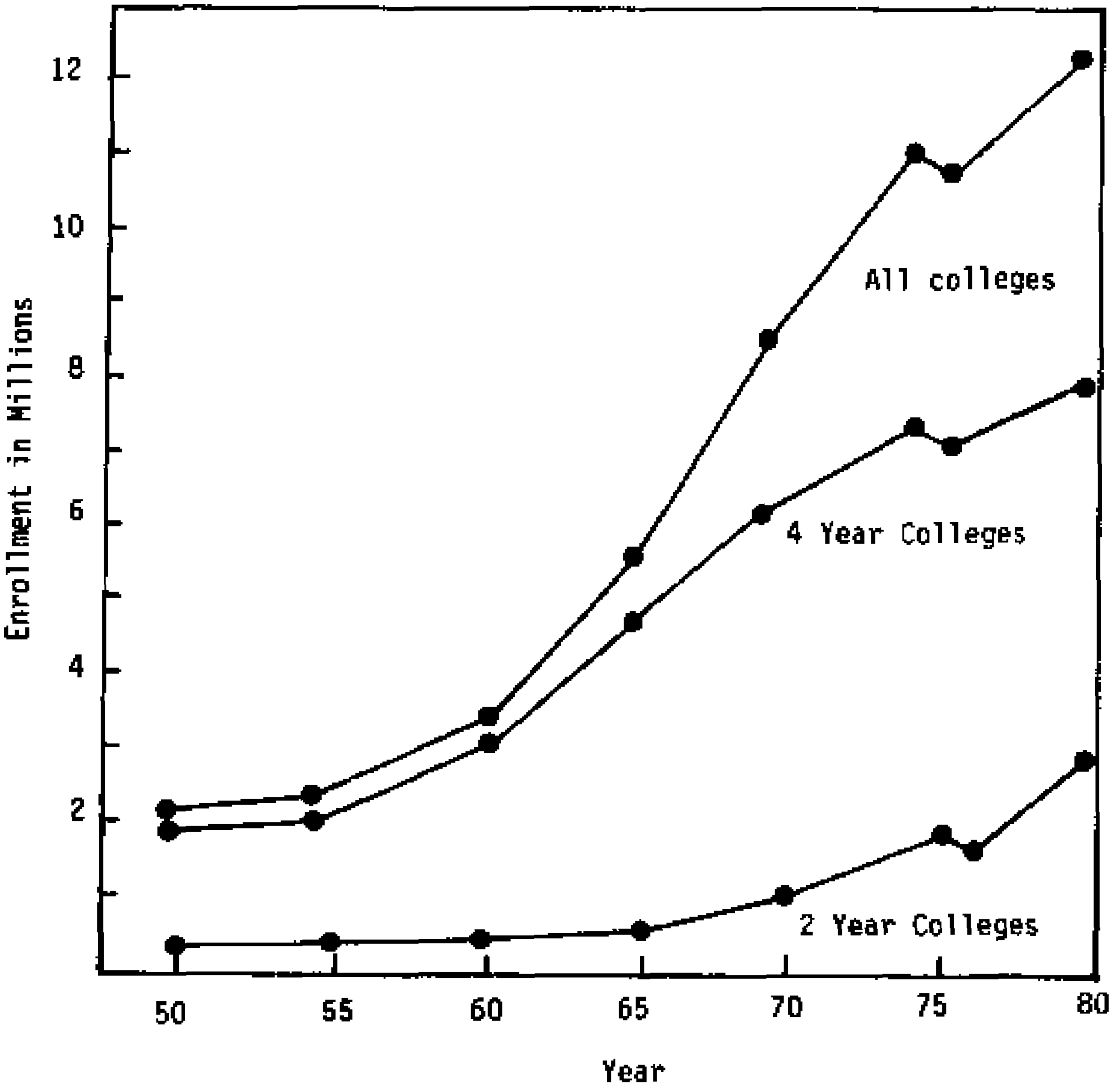


Figure II: Average Annual Headcount Enrollment

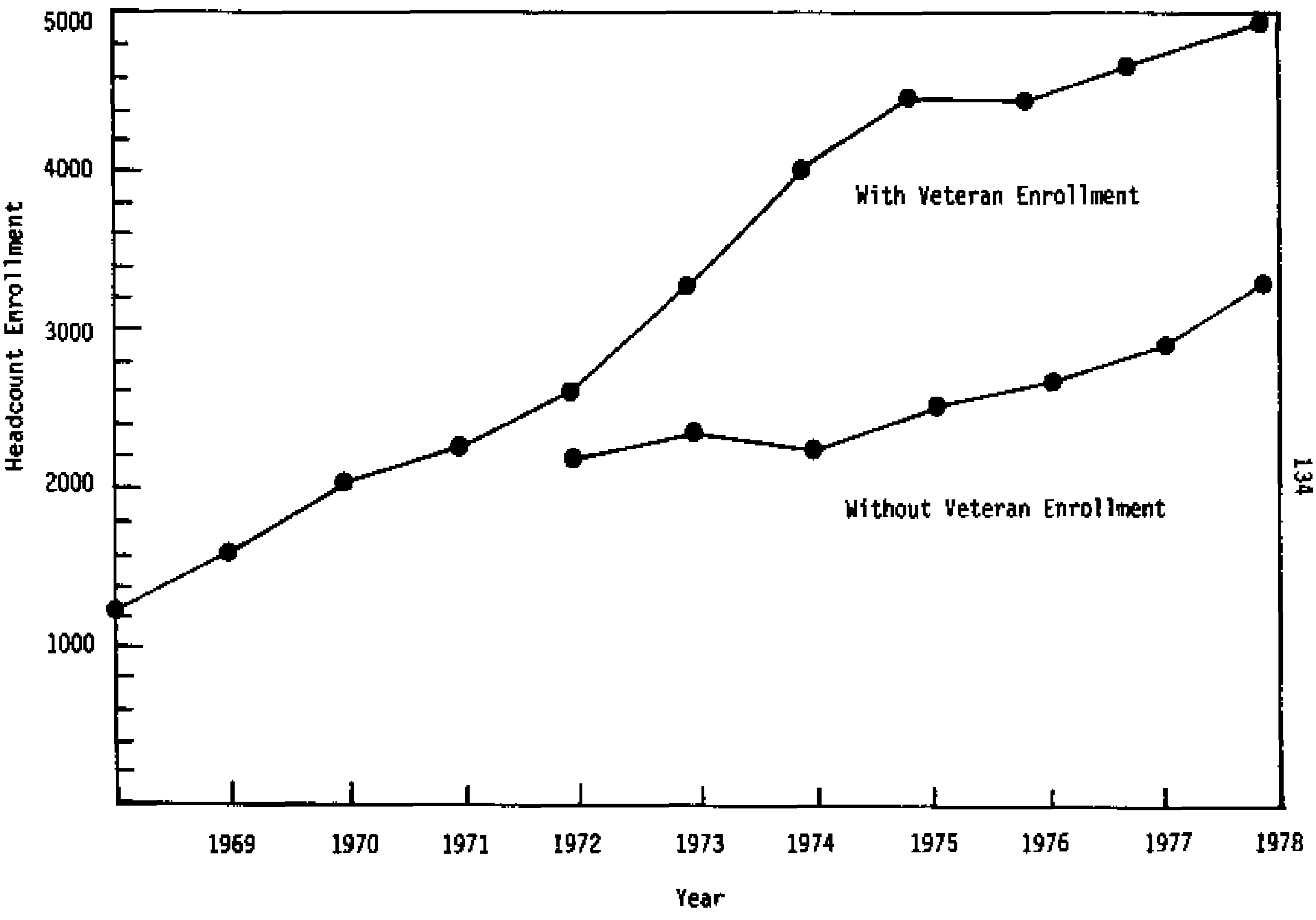


Figure III: Annual First-Time Student Enrollment and Annual Returning Student Enrollment

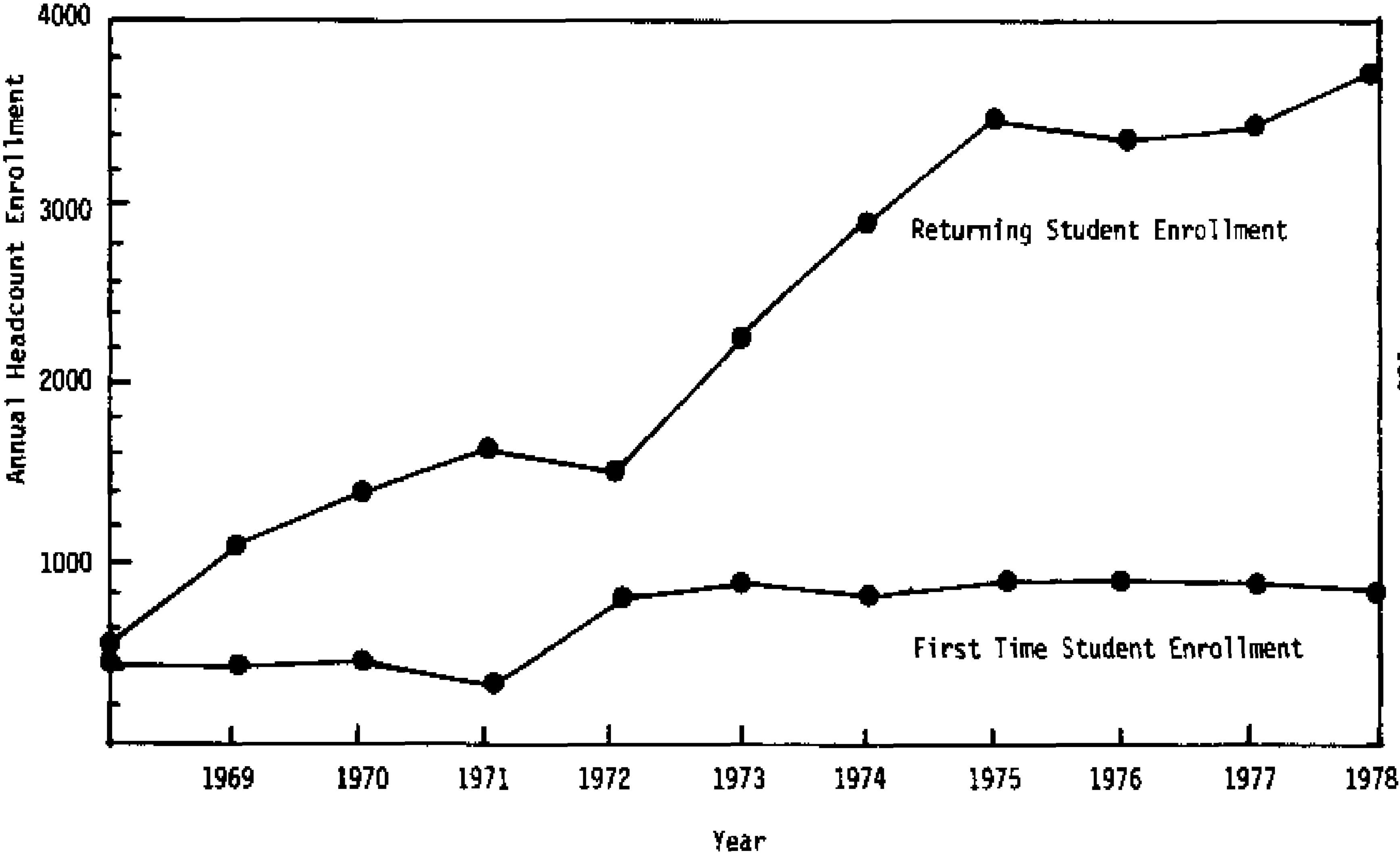


Figure IV: Total FTE Enrollment Per Quarter

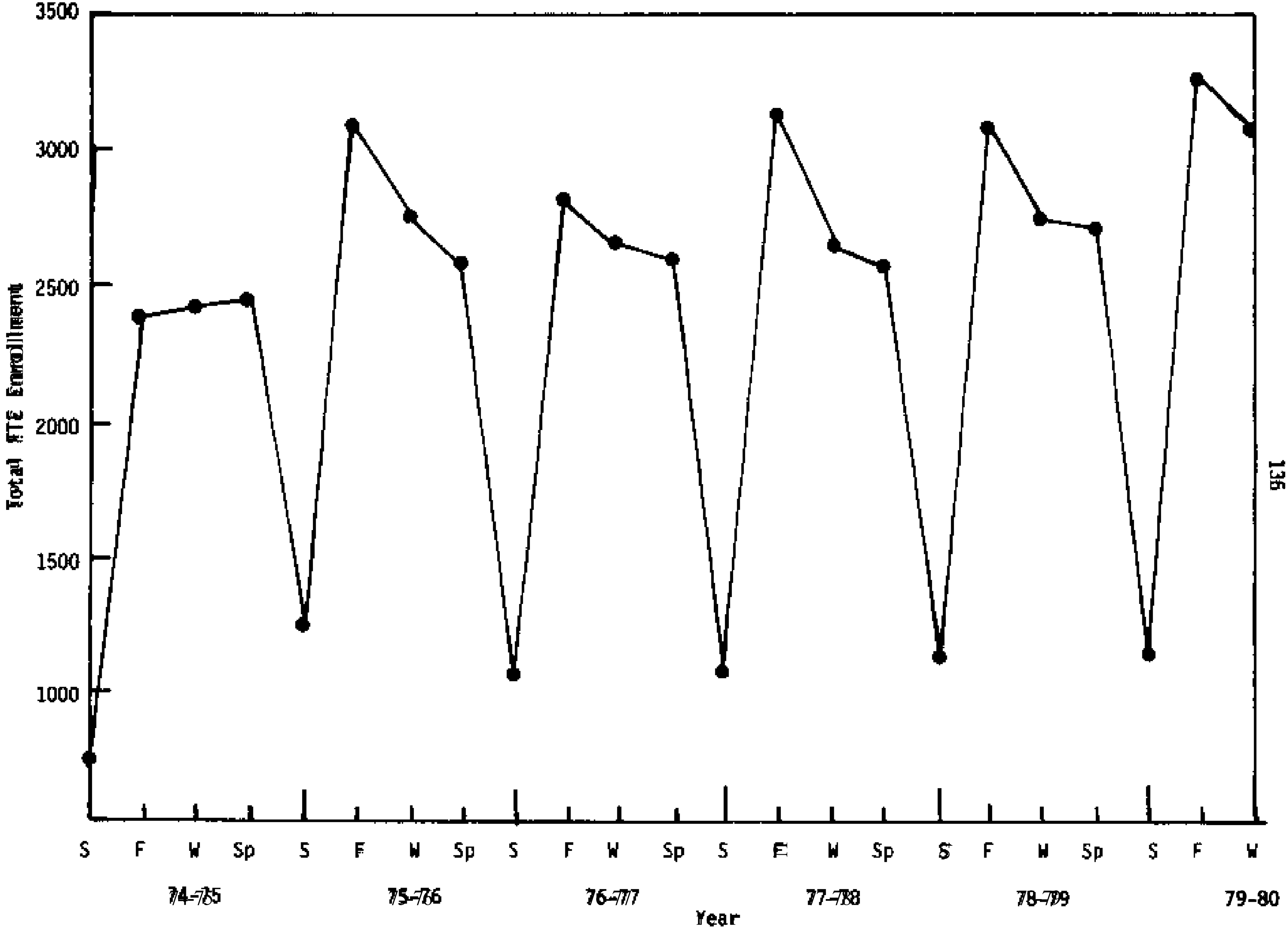
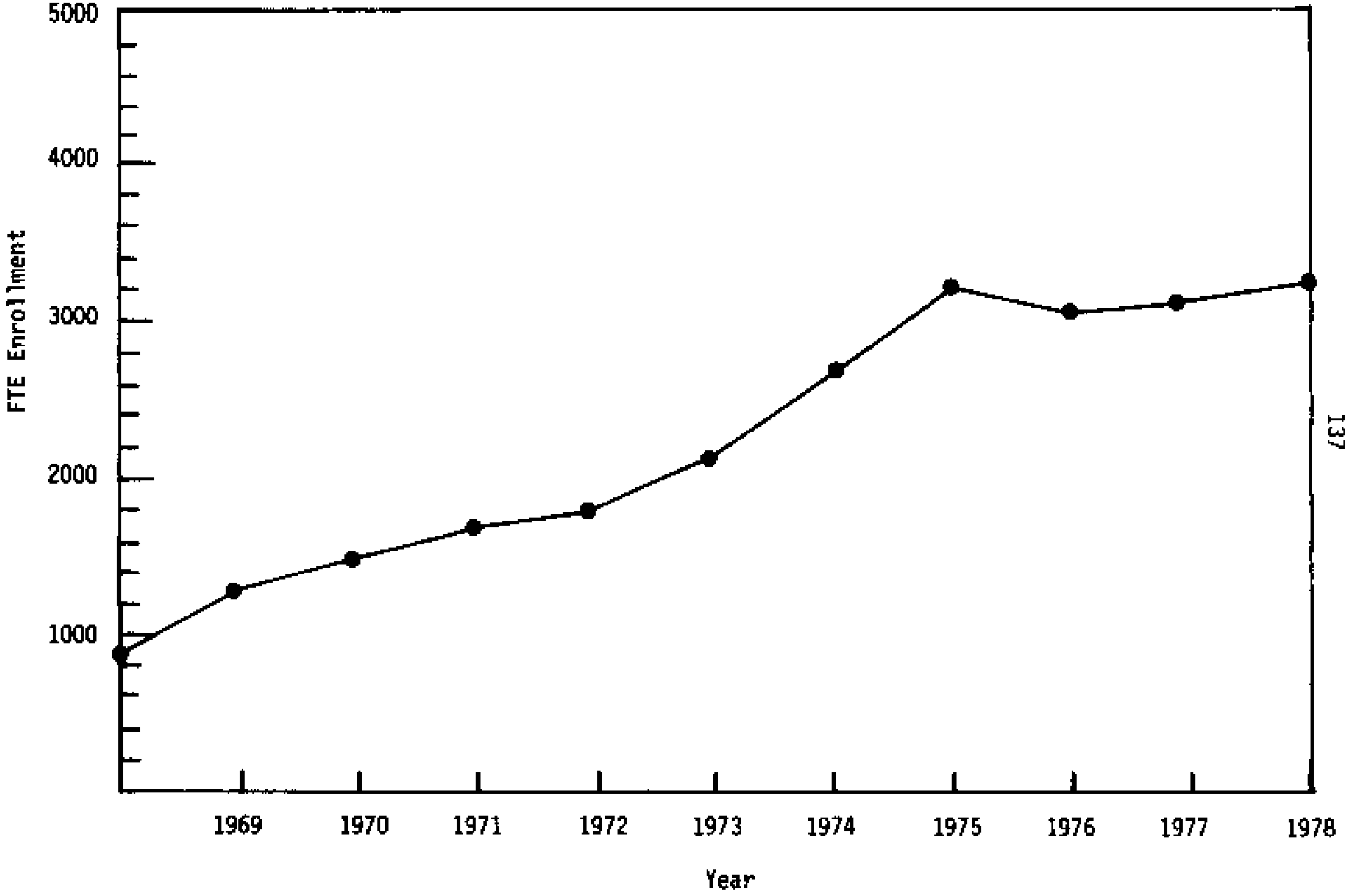


Figure V: Annual FTE Enrollment



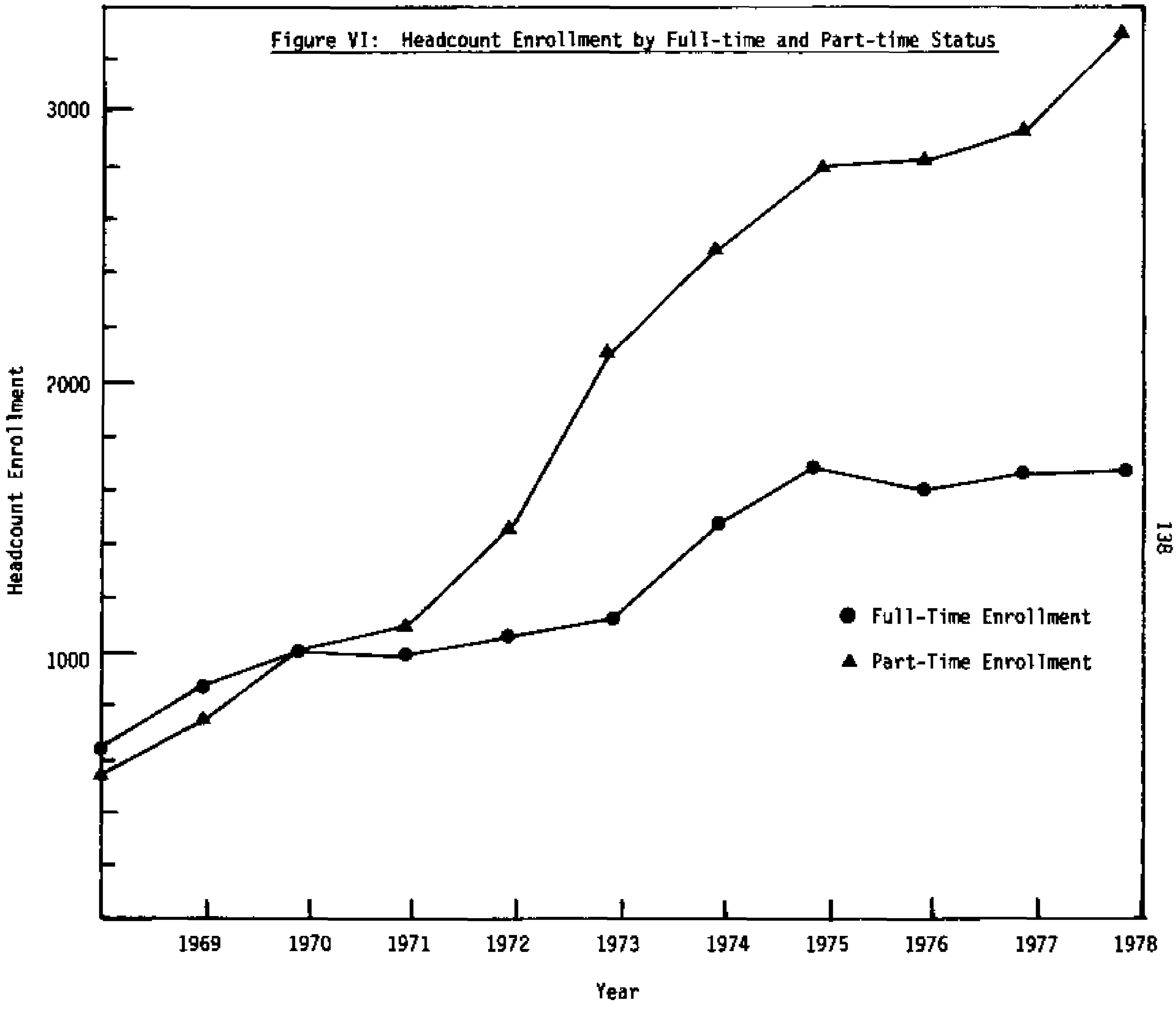
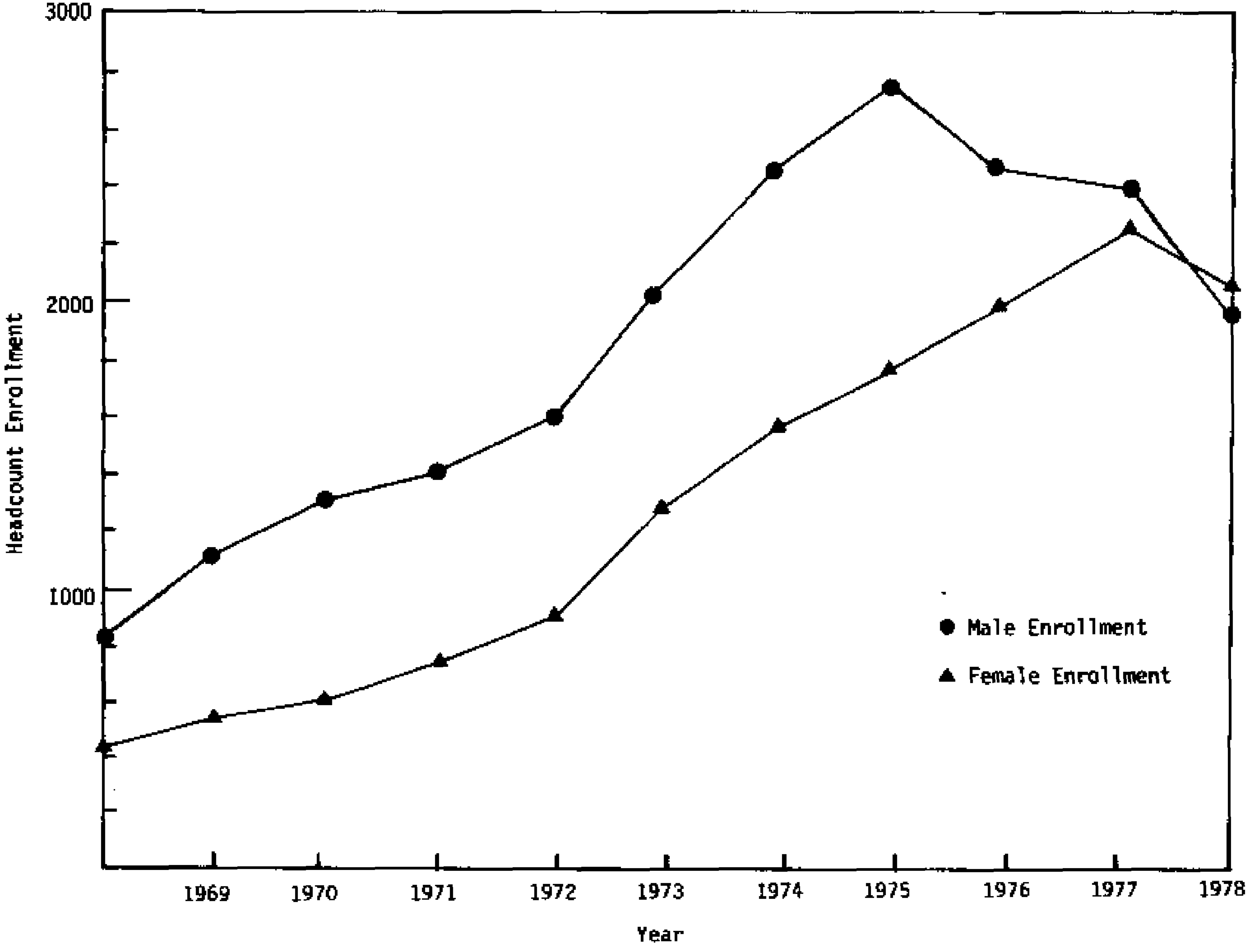


Figure VI: Headcount Enrollment by Full-time and Part-time Status

Figure VII: Headcount Enrollment by Sex



APPENDIX II

SIMPLE CORRELATIONS
BETWEEN
THE DEPENDENT AND POSSIBLE INDEPENDENT VARIABLES

The tables in Appendix II are the simple correlations between the dependent and independent variables selected for study, and were used to establish if relationships existed between the variables. In some cases the independent and dependent variables were shifted forward in time to determine if past events affected present enrollments. The time shifts were in quarters or years. The significance level for the correlations is given at the bottom of each table. Detailed discussion of the following tables is found in chapter 3.

TABLE IA: FACTORS AFFECTING AVERAGE ANNUAL HEADCOUNT ENROLLMENT

Independent Variables	Correlation Coefficients*	
	Not Including 1978-79 Data	Including 1978-79 Data
1. First-time student enrollment	0.7132	0.4592
2. Returning student enrollment	0.9979	0.9975
3. Part-time student enrollment	0.9977	0.9977
4. Full-time student enrollment	0.9850	0.9836
5. Male enrollment	0.9838	------
6. Female enrollment	0.9835	------
7. Veteran enrollment	0.9556	0.8191
8. High school graduates in S.A.	0.8013	0.8359
9. Total population in S.A.	0.9142	0.9310
10. Population, age 18-24 yrs, in S.A.	-0.9567	-0.8878
11. National two year college enrollment	0.9916	0.9931
12. Composite coincident economic indicator	0.8110	0.8406
13. Composite leading economic indicator	0.7243	0.7829
14. Local unemployment index	0.7533	0.7847
15. Time	0.9799	0.9813

*Significance level for all variables reported is less than or equal to 0.05

TABLE IIA: FACTORS AFFECTING AVERAGE HEADCOUNT ENROLLMENT SHIFTED FORWARD 1, 2, 3, and 4 YEARS

Factors	Correlation Coefficients*
1. Average annual headcount enrollment	
a. 1 year	0.9795
b. 2 years	0.9298
c. 3 years	0.8762
d. 4 years	0.8019
2. First-time student enrollment	
a. 1 year	0.9168
b. 2 years	0.8998
c. 3 years	0.7763
d. 4 years	0.4873
3. Part-time student enrollment	
a. 1 year	0.9874
b. 2 years	0.9437
c. 3 years	0.8757
d. 4 years	0.8308
4. Full-time student enrollment	
a. 1 year	0.9392
b. 2 years	0.8714
c. 3 years	0.8490
d. 4 years	0.9072
5. Composite national coincident economic indicator	
a. 2 years	0.9209
b. 3 years	0.8875
c. 4 years	0.9262
6. Composite national leading economic indicator	
a. 2 years	0.9109
b. 3 years	0.8684
c. 4 years	0.8980

*Significance level for all variables is less than or equal to 0.05

TABLE 18: FACTORS AFFECTING FALL HEADCOUNT ENROLLMENT

Independent Variables	Correlation Coefficients*	
	Not including 1978-79 Data	Including 1978-79 Data
1. First-time student enrollment	0.8529	0.8389
2. Returning student enrollment	0.9920	0.9934
3. Part-time student enrollment	0.9894	0.9909
4. Full-time student enrollment	0.9880	0.9866
5. Male enrollment	0.9253	------
6. Female enrollment	0.9834	------
7. Veteran enrollment	0.9153	0.8641
8. High school graduates in S.A.	0.7884	0.8250
9. Total population in S.A.	0.9236	0.9385
10. Population, 18-24 yrs., in S.A.	-0.9379	-0.8728
11. National 2 year college enrollment	0.9945	0.9953
12. CEI	0.7920	0.8260
13. LEI	0.7107	0.7715
14. Local unemployment index	0.7896	0.8154
15. Time	0.9814	0.9820

*Significance level for all variables is less than or equal to 0.05

TABLE IIB: FACTORS AFFECTING FALL HEADCOUNT ENROLLMENT SHIFTED FORWARD 1, 2, 3, and 4 YEARS

Factors	Correlation Coefficient*
1. Average Annual Headcount Enrollment	
a. 1 year	0.9813
b. 2 years	0.9462
c. 3 years	0.8916
d. 4 years	0.8660
2. First-time student enrollment	
a. 1 year	0.8928
b. 2 years	0.9113
c. 3 years	0.8174
d. 4 years	0.5877
3. Part-time student enrollment	
a. 1 year	0.9867
b. 2 years	0.9606
c. 3 years	0.8925
d. 4 years	0.8476
4. Full-time student enrollment	
a. 1 year	0.8913
b. 2 years	0.8862
c. 3 years	0.8607
d. 4 years	0.8739
5. Composite national coincident economic indicators	
a. 2 years	0.9136
b. 3 years	0.8964
c. 4 years	0.9317
6. Composite national leading economic indicators	
a. 2 years	0.8956
b. 3 years	0.8698
c. 4 years	0.9143

*Significance level for all variables is less than or equal to 0.05

TABLE IC: FACTORS AFFECTING RETURNING STUDENT ENROLLMENT

Independent Variables	Correlation Coefficients*
1. Part-time student enrollment	0.9853
2. Full-time student enrollment	0.9919
3. Male enrollment	0.9647
4. Female enrollment	0.9679
5. Veteran enrollment	0.9775
6. Total population in S.A	0.9112
7. Population, age 18-24 years, in S.A.	-0.9822
8. CEI	0.6203
9. LEI	0.4970
10. Local unemployment index	0.8094
11. Time	0.9517
12. National two year college enrollment	0.9826

---

*Significance level for all variables reported is less than or equal to 0.05

TABLE IIC: FACTORS AFFECTING RETURNING STUDENT ENROLLMENT SHIFTED FORWARD 1, 2, 3, AND 4 YEARS

Factors	Correlation Coefficients*
1. Average annual headcount enrollment	
a. 1 year	0.9640
b. 2 years	0.8734
c. 3 years	0.8133
d. 4 years	0.8286
2. First-time student enrollment	
a. 1 year	0.8413
b. 2 years	0.9130
c. 3 years	0.8337
d. 4 years	0.5430
3. Part-time student enrollment	
a. 1 year	0.9774
b. 2 years	0.9080
c. 3 years	0.8310
d. 4 years	0.7996
4. Full-time student enrollment	
a. 1 year	0.8994
b. 2 years	0.7511
c. 3 years	0.7347
d. 4 years	0.8674
5. Composite national coincident economic indicator	
a. 2 years	0.9402
b. 3 years	0.8915
c. 4 years	0.9150
6. Composite national leading economic indicator	
a. 2 years	0.9240
b. 3 years	0.8808
c. 4 years	0.8909

*Significance level for all variables reported is ≤ 0.05.

TABLE ID: FACTORS AFFECTING FIRST-TIME STUDENT ENROLLMENT

Factors	Correlation Coefficients*
1. Population, age 18-24 years in S. A.	-0.8045
2. Total population in S. A.	0.7493
3. Full-time student enrollment	0.7994
4. Part-time student enrollment	0.8699
5. Male enrollment	0.8601
6. Female enrollment	0.8575
7. National 2-year college enrollment	0.8442
8. High school graduates in S. A.	0.3176
9. CEI	0.7660
10. LEI	0.7695
11. Local unemployment index	0.4362
12. Time	0.8236
13. Veteran enrollment	0.6002

*Significance level ≤0.3. Most were at 0.2.

TABLE IID: ECONOMIC FACTORS AFFECTING FIRST-TIME STUDENT ENROLLMENT SHIFTED FORWARD 2, 3, AND 4 YEARS

Factors	Correlation Coefficients*
1. Composite national coincident economic indicators	
a. 2 years	0.7311
b. 3 years	0.7949
c. 4 years	0.8001
2. Composite national leading economic indicators	
a. 2 years	0.7144
b. 3 years	0.7535
c. 4 years	0.7903

*Significance level = 0.3.

TABLE IE: FACTORS AFFECTING TOTAL CREDITS OR FTE ENROLLMENT*

Factors	Correlation Coefficients**
1. First-time student enrollment	0.8295
2. Returning student enrollment	0.9946
3. Average annual headcount enrollment	0.9938
4. Fall headcount enrollment	0.9913
5. Veteran enrollment	0.9779
6. Total population in S. A.	0.8989
7. Full-time student enrollment	0.9942
8. Part-time student enrollment	0.9816
9. Female enrollment	0.9716
10. Male enrollment	0.9724
11. High school graduates	0.8610
12. CEI	0.7821***
13. LEI	0.7326***
14. Local unemployment index	0.8344****
15. Population, age 18-24 years, in S. A.	-0.9968
16. Time	0.9445

*FTE enrollment = total credits/45

Significance level for all variables except * ≤0.05

***Significance level for economic factors = 0.3

****Significance level = 0.02

TABLE IIE: FACTORS AFFECTING TOTAL CREDITS SHIFTED FORWARD 1, 2, 3, AND 4 YEARS

Factors	Correlation Coefficients*
1. Average annual headcount enrollment	
a. 1 year	0.9553
b. 2 years	0.8721
c. 3 years	0.8198
d. 4 years	0.9226
2. First-time student enrollment	
a. 1 year	0.7668
b. 2 years	0.9238
c. 3 years	0.7853
d. 4 years	0.2474
3. Part-time student enrollment	
a. 1 year	0.9656
b. 2 years	0.9148
c. 3 years	0.8075
d. 4 years	0.8867
4. Full-time student enrollment	
a. 1 year	0.8754
b. 2 years	0.6762
c. 3 years	0.8432
d. 4 years	0.9201
5. Composite national coincident economic indicator	
a. 2 years	0.9817
b. 3 years	0.7007
c. 4 years	0.5740
6. Composite national leading economic indicator	
a. 2 years	0.9400
b. 3 years	0.8573
c. 4 years	0.6014

*Significance level is ≤ 0.1.

TABLE IF: FACTORS AFFECTING PART-TIME STUDENT ENROLLMENT

Factors	Correlation Coefficients*
1. Veteran enrollment	0.9666
2. Total population in S. A.	0.9224
3. High school graduates in S. A.	0.9362
4. CEI	0.7580
5. LEI	0.6394
6. Local unemployment index	0.7129

*Significance level is ≤ 0.05.

TABLE IIF: ECONOMIC FACTORS AFFECTING PART-TIME STUDENT ENROLLMENT SHIFTED FORWARD 1 AND 2 YEARS

Factors	Correlation Coefficients*
1. Composite national coincident economic indicator	
a. 1 year	0.8255
b. 2 years	0.8066
2. Composite national leading economic indicator	
a. 1 year	0.7585
b. 2 years	0.7863

*Significance level is ≤ 0.05

TABLE IG: FACTORS AFFECTING FULL-TIME STUDENT ENROLLMENT

Factors	Correlation Coefficients*
1. Veteran enrollment	0.9730
2. Total population in S. A.	0.8638
3. High school graduates in S. A.	0.9571
4. CEI	0.5647
5. LEI	0.4252
6. Local unemployment index	0.8345

*Significance level is ≤ 0.01 except for factors 4 and 5 which were not significant

TABLE IIG: ECONOMIC FACTORS AFFECTING FULL-TIME STUDENT ENROLLMENT SHIFTED FORWARD 1 AND 2 YEARS

Factors	Correlation Coefficients*
1. Composite national coincident economic indicator	
a. 1 year	0.8205
b. 2 years	0.8883
2. Composite national leading economic indicator	
a. 1 year	0.6504
b. 2 years	0.6505

*Significance level is ≤ 0.05

TABLE IH: FACTORS AFFECTING MALE ENROLLMENT

Factors	Correlation Coefficients*
1. Veteran enrollment	0.9954
2. Total population in S. A.	0.8063
3. High school graduates in S. A.	0.9891
4. CEI	0.6170
5. LEI	0.4595
6. Local unemployment index	0.6978

*Significance level is ≤ 0.01 except for factor 5 which is not significan

TABLE IIH: ECONOMIC FACTORS AFFECTING MALE ENROLLMENT SHIFTED FORWARD 1 AND 2 YEARS

Factors	Correlation Coefficients*
1. Composite national coincident economic indicator	
a. 1 year	0.8728
b. 2 years	0.8833
2. Composite national leading economic indicator	
a. 1 year	0.7391
b. 2 years	0.9067

*Significance level is ≤ 0.05

TABLE II: FACTORS AFFECTING FEMALE ENROLLMENT

Factors	Correlation Coefficients*
1. Total population in S. A.	0.8603
2. High school graduates in S. A.	0.8754
3. CEI	0.6170
4. LEI	0.4595
5. Local unemployment index	0.6978

*Significance level is ≤ 0.01 except for factor 4 which is not significan

TABLE III: ECONOMIC FACTORS AFFECTING FEMALE ENROLLMENT SHIFTED FORWARD 1 AND 2 YEARS

Factors	Correlation Coefficients*
1. Composite national coincident economic indicator	
a. 1 year	0.7605
b. 2 years	0.7610
2. Composite national leading economic indicator	
a. 1 year	0.7008
b. 2 years	0.6936

*Significance level is ≤ 0.1

TABLE IJ: FACTORS AFFECTING VETERAN ENROLLMENT

Factors	Correlation Coefficients*
1. Total population in S. A.	0.8086
2. Population, age 18-24 years in S. A.	0.9639
3. National 2-year college enrollment	0.9598
4. Time	0.9062
5. CEI	0.5951
6. LEI	0.4339
5. Local unemployment index	0.7436

*Significance level is ≤ 0.05

TABLE IIJ: ECONOMIC FACTORS AFFECTING VETERAN ENROLLMENT LAGGED 1, 2, 3, AND 4 YEARS

Factors	Correlation Coefficients*
1. Composite national coincident economic indicator	
a. 1 year	0.8566
b. 2 years	0.8979
c. 3 years	0.5770
d. 4 years	0.5357
2. Composite national leading economic indicator	
a. 1 year	0.7036
b. 2 years	0.9131
c. 3 years	0.7511
d. 4 years	0.4952

*Significance level is ≤ 0.5 except for both factors lagged 3 and 4 years which are insignificant

APPENDIX III

ENROLLMENT FORECASTS

The tables in Appendix III are the forecasts produced for the dependent variables selected for study. The independent variables used in the forecasting model are listed along with the relevant statistics. Unless specified in the tables, the forecasting model was multiple regression. In several instances, the independent variables were shifted forward for several time shifts (in quarters or years) in the same equations. When a variable is shifted forward for more than one time shift in the same equation, the variable is listed once with the various time shifts in parentheses. CEI's (t+6, 7, 8, 9), therefore, indicates that the coincident economic indicator was used four times in the same equation with forward time shifts of 6, 7, 8, and 9 quarters. The tables in Appendix III are thoroughly discussed in chapters 3 and 4.

TABLE A: QUARTERLY FORECASTS FOR HEADCOUNT ENROLLMENT

FORECAST NUMBER	INDEPENDENT VARIABLES	MULTIPLE R	MSE	STD. DEV.	$\overline{e}$	PERCENT ERROR IN FORECAST				
						SUMMER	FALL	WINTER	SPRING	YEAR
1. (a) (b)	T; LEI's (t+4), (t+8), CEI's (t+5), (t+6), (t+7), (t+8)	0.9884 0.9921	1,463 14,940	32.9 134.5	4.8 5.9	2.4 8.1	0.0008 0.4	3.5 8.7	2.1 18.5	78-79 79-80
2.	T;CEI's (t+4, 5, 6, 7, 8, 9, 10); LEI's (t+4, 5, 6, 7, 8, 9, 10)	0.9803	16,490	126.4	22.9	1.4	8.2	3.4	----	79-80
3.	T; CEI's (t+6, 7, 8, 9); LEI (t+8)	0.9913	22,397	149.6	1.4	6.7	2.2	5.4	----	79-80
4.	T; CEI's (t+5, 7, 8); LEI's (t+4, 8)	0.9923	19,066	138.1	2.6	8.6	1.5	8.1	----	79-80
5.	T; CEI (t+10)	0.9798	52,350	222.5	53.4	0.8	3.1	4.2	----	79-80
6.	Time-Decomposition	0.9956	11,041	104.9	5.3	1.4	6.9	8.2	----	79-80

TABLE AI: FORECASTS AND ACTUAL ENROLLMENT BY QUARTER

Quarter	Actual	1a	1b	2	3	4	5
Summer 1978	2374	2316	2348	2401	2395	2357	2391
Fall 1978	5312	5316	5331	5446	5441	5357	5377
Winter 1979	4728	4895	5004	4927	4968	4996	4839
Spring 1979	4918	4815	4894	4723	4887	4927	4807
Summer 1979	2579	2713	2807	2617	2751	2802	2600
Fall 1979	6045	5831	6069	5548	5914	6133	5860
Winter 1980	5517	5945	5693	5329	5817	5963	5287
Spring 1980				5286	5474	5991	5279
Summer 1980					3152		2823
Fall 1980					6868		6386
Winter 1981							5768
Spring 1981							5766
Summer 1981							3088
Fall 1981							6826

TABLE B: FALL ENROLLMENT FORECASTS USING ANNUAL DATA

Forecast Number	Independent Variables	Multiple R	MSE	Std. Dev.	$\bar{e}$	Percent Error in Forecasts	
						78-79	79-80
1.	Hd. ct. E (t+1), LEI (t+2)	0.9934	137,514	366.0	23.4	5.9	8.6
2.	Time: LEI (t+2)	0.9927	9,940	99.7	0.0	4.1	1.9
3.	LEI (t+2), (t+4)	0.9922	30,261	174.0	0.0	1.5	12.9
4.	Time; LEI (t+3)	0.9817	60,389	251.7	0.0	5.6	0.4
5.	Time; CEI (t+3)	0.9852	51,158	226.2	0.0	18.7	0.6
6.	Time	0.9814	64,342	253.7	0.0	7.8	0.3
7.	Part-time (t+1)	0.9635	55,395	235.4	0.0	2.0	5.5
8.	Hd. ct. (t+1)	0.9605	64,490	254.0	0.0	2.5	8.5
9.	First-Time (t+2)	0.9070	137,179	370.4	0.3	9.5	17.8
10.	Time (exponential)	0.9686	203,140	450.7	3.4	16.7	20.4
11.	Time (power)	0.9794	95,177	306.6	34.5	2.0	8.9
12.	Time (logarithmic)	0.9342	246,619	496.6	0.1	7.6	16.2
13.	Part-time (t+1), LEI (t+2)	0.9854	41,620	204.4	0.2	0.4	1.2

TABLE BI: FALL HEADCOUNT ENROLLMENT FORECASTS AND ACTUAL ENROLLMENT - ANNUAL DATA USED

YEAR	ACTUAL	FORECASTS		
		4	5	6
1968-69	1236	1257	1255	1270
1969-70	1832	1687	1702	1703
1970-71	2220	2104	2135	2135
1971-72	2305	2596	2575	2567
1972-73	2676	3022	3014	3000
1973-74	3399	3328	3429	3432
1974-75	4034	3859	3852	3864
1975-76	4959	4421	4303	4296
1976-77	4616	4914	4756	4729
1977-78	5161	5146	5170	5161
1978-79	5312	5416	5558	5593
1979-80	6045	6022	6008	6025
1980-81	----	6526	6455	6458
1981-82	----	6996	6906	6890

TABLE C: ANNUAL FTE FORECASTS

Forecast Number	Independent Variables	Multiple R	MSE	Std. Dev.	$\bar{e}$	Percent Error in Forecasts 78-79
1.	Hd. ct. (t+1) (power)	0.9915	14,697	121.2	2.0	3.0
2.	Hd. ct. (t+1) (exponential)	0.9740	390,152	624.6	1.4	9.0
3.	Hd. ct. (t+1) (logarithmic)	0.9826	18,056	424.9	0.1	0.8
4.	CEI (t+2), CEI (t+4)	0.9817	25,078	158.4	0.1	5.6
5.	CEI (t+2), CEI (t+4), T	0.9941	7,559	86.9	0.0	13.0
6.	T, LEI (t+2), (t+3)	0.9817	5,406	73.5	0.2	8.8
7.	T, LEI (t+2), (t+4)	0.9953	6,459	80.4	0.1	3.0
8.	Part-time (t+1), (t+4)	0.9963	2,196	46.9	0.0	53.4
9.	Part-time (t+1)	0.9880	12,535	111.9	2.2	4.1
10.	Part-time (t+1) (exponential)	0.9783	28,552	169.0	0.9	9.7
11.	Part-time (t+1) (power)	0.9949	9,474	97.3	0.6	2.0
12.	Part-time (t+1) (logarithmic)	0.9843	16,272	127.6	0.0	1.4
13.	Part-time (t+2)	0.9426	48,998	221.4	0.1	0.2
14.	Part-time (t+2) (exponential)	0.9307	77,225	277.9	8.2	9.9
15.	Part-time (t+2) (power)	0.9705	34,320	185.2	4.5	5.5
16.	Part-time (t+2) (logarithmic)	0.9695	26,466	162.7	0.1	2.1
17.	P.T./F.T. (t+1) (logarithmic)	0.9445	56,458	237.6	0.1	5.7
18.	P.T./F.T. (t+1) (exponential)	0.9576	55,922	236.2	11.8	8.5
19.	Part-time (t+1), CEI (t+2)	0.9919	848	29.1	0.0	3.6
20.	P.T./F.T. (t+1) (power)	0.9574	33,073	181.8	5.9	4.3
21.	Part-time (t+2), LEI (t+2)	0.9958	3,648	60.4	0.0	9.0
22.	Hd. ct. (t+1), LEI (t+2), CEI (t+3)	0.9999	29,508	171.8	3.1	5.5
23.	CEI (t+2), Full-time (t+3), (t+2)	0.9999	9.8	3.1	0.8	32.8
24.	CEI (t+2), Full-time (t+3)	0.9924	5,619	75.0	0.0	20.7
25.	CEI (t+3), Full-time (t+3)	0.8498	103,480	321.7	0.0	0.3
26.	T, CEI (t+4)	0.9910	11,690	108.1	0.0	16.9
27.	T	0.9766	30,159	173.7	0.0	12.9

TABLE C: ANNUAL FTE FORECASTS (Concluded)

Forecast Number	Independent Variables	Multiple R	MSE	Std. Dev.	$\overline{e}$	Percent Error in Forecasts 78-79
28.	Time (exponential)	0.9688	64,859	254.7	0.0	32.8
29.	Time (logarithmic)	0.9325	85,030	291.6	0.1	4.1
30.	Time (power)	0.9793	40,071	199.6	15.4	3.0
31.	T, CEI (t+3)	0.9805	25,470	159.6	0.3	19.6
32.	Part-time (t+1), (t+4)	0.9792	11,948	109.3	0.1	2.7
33.	Hd. ct. (t+1)	0.9607	22,346	149.5	0.0	2.8
34.	Fall Hd. ct. (t+1)	0.9659	35,083	187.3	0.1	8.4
35.	Fall Hd. ct. (t+1) (exponential)	0.9226	67,191	259.1	4.0	12.1
36.	Fall Hd. ct. (t+1) (logarithmic)	0.9293	36,973	192.3	0.1	1.5
37.	Fall Hd. ct. (t+1) (power)	0.9788	31,168	176.4	6.0	6.7
38.	T, CEI (t+2)	0.9911	15,564	124.8	0.0	9.2
39.	P.T. (t+1), (t+2); LEI (t+2)	0.9999	91,664	302.8	0.0	6.7

TABLE D: QUARTERLY FTE FORECASTS

FORECAST NUMBER	INDEPENDENT VARIABLES	MULTIPLE R	MSE	STD. DEV.	$\bar{e}$	PERCENT ERROR IN 1979 FORECAST		
						SUMMER	FALL	WINTER
1.	T, CEI (t+6, 7, 8, 9); LEI (t+8)	0.9871	14,340	119.7	0.0	14.2	0.8	
2.	T	0.9655	19,541	139.5	8.9	5.4	1.3	
3.	T, LEI (t+2) Time Dec. Subtraction	0.9966	3,543	59.3	4.7	3.6	3.1	6.6
4.	Same as 3 - (Division Method)	0.9965	3,803	61.1	8.1	3.6	3.5	6.5
5.	T, CEI (t+6) (Time Decom.) Sub.	0.9963	3,903	62.0	7.7	4.0	2.1	5.1
6.	Same as 5 - Division	0.9962	3,901	62.0	7.6	3.9	2.2	5.2
7.	T, Local unemp. (t+7) (Time Decom.) Subtraction	0.9966	3,839	61.4	8.3	0.8	7.3	10.3
8.	Same as 7 - Division	0.9966	3,839	61.4	8.3	0.8	7.3	10.3
9.	T, CEI (t+6), Local Unemp.(t+7) (Time Decom.) Subtraction	0.9966	2,456	48.7	9.3	3.8	3.3	6.0
10.	T, CEI (t+6); LEI (t+7) (Time Decom.) Subtraction	0.9964	3,832	61.8	3.8	2.6	5.6	

TABLE E: RETURNING STUDENT ENROLLMENT FORECASTS USING ANNUAL DATA

FORECAST NUMBER	INDEPENDENT VARIABLES	MULTIPLE R	MSE	STD. DEV.	$\bar{e}$	PERCENT ERROR IN 78-79 FORECAST
1.	Time, LEI (t+2)	0.9997	358.2	18.9	1.0	11.9
2.	LEI (t+2), (t+3), T	0.9932	13,805	117.5	1.0	6.4
3.	CEI (t+2), (t+4)	0.9737	52,881	230.0	19.0	1.0
4.	Part-time (t+2); LEI (t+2)	0.9728	28,219	168.0	5.0	1.2
5.	First-time (t+2)	0.9109	87,616	296.1	0.0	9.9
6.	Hd. ct. (t+1); CEI (t+3)	0.9559	6,579	81.1	1.8	12.5
7.	Part-time (t+1); LEI (t+2); Part-time (t+3)	0.9999	89.6	9.5	0.6	7.0

TABLE F: QUARTERLY RETURNING STUDENT ENROLLMENT FORECASTS

FORECAST NUMBER	INDEPENDENT VARIABLES	MULTIPLE R	MSE	STD. DEV.	$\bar{e}$	PERCENT ERROR IN 78-79 FORECASTS	
						SUMMER	WINTER
1.	T, CEI (t+5, 6, 7, 8, 9) LEI (t+8)	0.9294	14,039	118.4	5.5	3.6	11.2
2.	T, CEI (t+10)	0.9902	12,332	110.0	15.9	4.3	1.2
3.	T, LEI (t+4, 8) CEI (t+5, 6, 7, 8)	0.9992	966	31.1	1.8	8.9	8.3

TABLE G: FIRST-TIME STUDENT ENROLLMENT FORECASTS

NUMBER	VARIABLES	R	MSE	DEV.	$\bar{e}$	PERCENT ERROR IN 78-79 FORECASTS
1.	LEI (t+4)	0.7851	19,127	138.3	0.1	7.4
2.	Time	0.8562	13,326	115.4	1.1	26.2
3.	CEI (t+4)	0.8113	17,020	130.5	10.2	17.4
4.	Moving average					7.1

TABLE H: AVERAGE ANNUAL HEADCOUNT ENROLLMENT FORECASTS

FORECAST NUMBER	INDEPENDENT VARIABLES	MULTIPLE R	MSE	STD. DEV.	$\overline{e}$	PERCENT ERROR IN 78-79 FORECAST
1.	T, LEI (t+2)	0.9943	17,082	130.7	0.0	6.7
2.	First-time (t+1.2)	0.9853	21,275	145.9	8.2	9.2
3.	T, LEI (t+2, 3)	0.9947	15,885	126.0	9.0	5.8
4.	T, CEI (t+2, 3); LEI (t+2, 3)	0.9987	3,887	62.3	5.7	20.0
5.	Part-time (t+1)	0.9735	30,509	174.7	11.0	2.9
6.	Hd. ct. (t+1); CEI (t+3)	0.9944	4,879	69.9	0.0	8.4
7.	First-time (t+4), CEI (t+2)	0.9988	7,646	81.0	32.0	17.8
8.	Hd. ct. (t+3, 4); LEI (t+2)	0.9999	92	1.7	0.0	16.3
9.	First-time (t+2)	0.8912	336,837	1,807.3	101.0	10.6

References

1. Ascher, William. *Forecasting: An Appraisal for Policy-Makers and Planners.* Baltimore, Md.: The John Hopkins University Press, 1978, Chpt. 4.

2. Banks, Jerry; and Hohenstein, C. L. *Conceptual Models and Procedures for Predicting Higher Education Enrollment.* Paper presented at Operations Research Society of America National Meeting. Detroit, Mich., 1970.

3. Bell, Daniel. *Toward the year 2000.* Boston, Mass.: Houghton Mifflin Co., 1968.

4. Bowen, Howard R. *Higher Education: A Growth Industry?* Educational Record, Summer, 1974, 147-158.

5. Bradley, Jr., A. Paul; and Lehmann, Timothy. *Attrition at a Non-traditional Institution.* Paper presented at The Association for Institutional Research, S. Louis, Missouri, 1975, 283-288.

6. The Carnegie Commission on Higher Education. *New Students and New Places.* New York: McGraw-Hill Book Co., 1971.

7. The Carnegie Commission on Higher Education. *Priorities for Action: Final Report of the Carnegie Commission on Higher Education.* New York: McGraw-Hill Book Co., 1973.

8. The Carnegie Commission on Higher Education. *Toward a Learning Society.* New York: McGraw-Hill Book Co., 1973.

9. The Carnegie Council Series. *More Than Survival.* San Francisco, Calif.: Jossey-Bass Pub., 1975.

10. Cartter, Allan M. *Higher Education Under Steady-State Conditions.* Paper presented at the Asso. for Institutional Research, Vancouver, British Columbia, 1973, 18-22.

11. Caspar, R. *Ein Modell zur Ermittlung Optimaler Zulassungszahlen Von Studienanfangern.* His-Briege, Heft Bl. 1969.

12. Centra, John A. *College Enrollment in the 1980s: Projections and Possibilities.* Journal of Higher Education, Jan/Feb, 1980, 19-39.

13. Cheit, Earl F. *The New Depression in Higher Education.* New York: McGraw-Hill Book Co., 1971.

14. Conner, Aikin. *Community College Enrollment Projections.* Washington, D.C. American Association of Junior Colleges, 1971, 1-14.

15. Dietze, H. D. *Studie zur Kapazitatsemittlung and Universitatem.* Germany: University of Saar, 1969.

16. Dresch, Stephen P. *Educational Saturation: A Demographic-Economic Model.* AAUP Bulletin, Autumn, 1975, 239-247.

17. Folger, John K. *On Enrollment Projections. Clearing up the Crystal Ball.* Journal of Higher Education, June, 1974, 405-414.

18. Froomkin, J. *Changing Credential Objectives of Students in Postsecondary Education.* Washington, D.C.: U.S. Department of Health, Education, and Welfare, Contract #0574257, 1974.

19. Gani, J. *Formulae for Projecting Enrollments and Degrees Awarded in Universities.* Journal of the Royal Statistical Society A 126, 1963, 400-409.

20. Gilchrist, Warren. *Statistical Forecasting.* London: John Wiley and Sons, 1976.

21. Glenny, Lyman A. *Demographic and Related Issues for Higher Education in the 1980s.* Journal of Higher Education, July/Aug, 1980, 363-378.

22. Graziano, Anthony F. *Possible Limiting Factors Operating on Junior College Growth.* Champaign, Ill.: Illinois Univ., 1972.

23. Harrington, Fred H. *The Future of Adult Education.* San Francisco, Calif.: Jossey-Bass Publisher, 1977.

24. Jellema, Wm. W. *The Numbers Game: A Study of Enrollment Patterns in Private Colleges and Universities.* Association of American Colleges, Washington, D. C., 1972.

25. Kibbee, Robert J. *The Hazards of Planning - Predicting Public Policy.* Paper presented at the Association of Institutional Research, Vancouver, British Columbia, 1973, 23-26.

26. Koenig, Herman E.; Keeney, M.G.; and Zemach, R. *A Systems Model for Management Planning and Resource Allocation in Institutions of Higher Education.* East Lansing, Michigan: Michigan State University, 1968.

27. Leslie, L. L.; and Miller, H. F., Jr. *Higher Education and the Steady State.* ERIC/Higher Education Research Report No. 4 Washington, D.C.: American Association for Higher Education, 1974.

28. Lightfield, E. Timothy. *They Do Come Back! Enrollment Tracking at the Community College.* Paper presented at the Association for Institutional Research, St. Louis, Missouri, 1975, 253-259.

29. Lovell, C. C. Student Flow Models: A Review and Conceptualization. Western Interstate Commission for Higher Education, Boulder, Colo. National Center for Higher Education Management Systems, 1971.

30. Mangelson, Wayne L.; Norris, Donald M.; Poulton, Nick L.; and Seely, John A. Projecting College and University Enrollments: Analyzing the Past and Focusing the Future. Center for the Study of Higher Education, School of Ed., U. of Mich., Ann Arbor, Mich., 1973.

31. Mangelson, Wayne L. and Others. National Enrollment Projection Studies. Planning for Higher Education. Vol. 3, No. 6:1/6, 1974.

32. Monroe, Charles R. Profiles of the Community College. San Francisco, Calif.: Jossey-Bass Pub., 1973.

33. National Bureau of Economic Research. Cyclical Indicators. Business Conditions Digest, March, 1980, 60, 97.

34. Nie, Norman H.; Hull, C. Hadlai; Jenkins, Jean G.; Steinbrenner, Karen; Bent, Dale H. SPSS Statistical Package for the Social Sciences. 2nd. Ed. New York: McGraw-Hill Book Co., 1975.

35. Newton, Robert D. Assessing the Impacts of Future Student Demand: An Application of a Demographically Differentiated Projection Model. Paper presented at the Association of Institutional Research, Los Angeles, Calif., 1976, 113-118.

36. Norris, Donald M. Enrollment Projection Strategies in an Uncertain Environment. Paper presented at the Association of Institutional Research, Los Angeles, Calif., 1976, 137-140.

37. Office of Institutional Research. TNCC Statistical Abstract 1977-1978. Thomas Nelson Community College, 1978.

38. Office of Institutional Research. Tentative Ten-Year Enrollment Projections. Fiscal Years 1979-1988. Montgomery College, Md., 1977.

39. Office of Program Planning and Fiscal Management, State of Washington. Higher Education Enrollment Projection Model. 1970.

40. Renssalaer Research Corporation. Construction and Analysis of a Prototype Planning Simulation for Projecting College Enrollments. Troy, New York, 1970.

41. Russell, C. Neil; and Hoffman, B. B. Three Years into the Future and Beyond with a Postsecondary Demand and Enrollment Model. Paper presented at the Assoc. for Institutional Research, Los Angeles, Calif., 1976, 109-112.

42. Salley, Charles D. Short-term Enrollment Forecasting for Accurate Budget Planning. Journal of Higher Education, Vol. 50. No.3, 1979, 323-333.

43. Smith, Wayne. A Student Flow Model. Mimeographed. Los Angeles, Calif.; Office of Advanced Planning, University of Calif., Los Angeles, 1970.

44. State Council of Higher Education for Virginia. Code of Virginia, Vol. 5, Title 23, 1977, 4.

45. State Council of Higher Education for Virginia. The Virginia Plan. Richmond: SCHEV, 1974.

46. Southern Regional Education Board. Fact Book on Higher Education in the South 1977 and 1978. Atlanta, Georgia: SREB, 1978.

47. Suslow, Sidney. Benefits of a Cohort Survival Projection Model. New Directions for Institutional Research, 13, Spring, 1977.

48. Thonstad, T. A Mathematical Model The Norwegian Education System. Norway: Organization for Economic Cooperation and Development, 1967, 19-42.

49. Tsai, Yung-Mei. Sociological Determinants of College Attendance: Some Causal Models. Boulder, Colorado: Western Interstate Commission for Higher Education, 1973.

50. U.S. Dept. of Commerce, Bureau of Census. Statistical Abstract of the U.S., 1973.

51. U.S. Dept. of Commerce, Bureau of Census. Statistical Abstract of the U.S., 1976.

52. U.S. Dept. of Health, Education, and Welfare, National Center for Education Statistics, Projections of Education Statistics to 1986-87. U.S. GPO, Washington, D.C., 1978.

53. Virginia Community College System. Policy Manual, Current, Section 9.

54. Virginia Community College System. Student Enrollment Booklet, 1971-79.

55. Virginia Dept. of Education. Facing Up. March 74-March 79.

56. Virginia Dept. of Planning and Budget. Budget Manual, Appendix M. 80-82 Biennium, 24-27.

57. Virginia Dept. of Planning and Budget. Population Projections. Virginia Counties and Cities, 1980-2000, June, 1979.

58. Virginia Employment Commission, Area Manpower Review, Hampton-Newport News, Aug., 1979.

59. Virginia State Legislature. Virginia Appropriation Act, 1978. Section 178, p. 166.

60. Wasik, John L. The Development of a Mathematical Model to Project Enrollments in a Community College System. North Carolina State University, 1971.

61. Wheelwright, Steven C.; and Makridalus, Spyros. Forecasting Methods for Management, 2nd. Ed. New York: John Wiley and Sons, 1977.

62. Wing, Paul. Higher Education Enrollment Forecasting. A Manual for State Level Agencies, NCHEMS, 1974.

VITA

Sue C. Lawrence

Address:

16 Carroll Drive

Poquoson, Virginia 23662

Birthplace:

Cincinnati, Ohio

Education:

1980	Ed.D	Higher Education Administration The College of William Mary Williamsburg, Virginia 23185
1955	M.A.	Teaching Physics Vanderbilt University Nashville, Tennessee
1954	B.A.	Physics and Mathematics Hollins College Hollins College, Virginia

Experience:

1968 - Present:
Thomas Nelson Community College, Hampton, Virginia
Professor of Physics and Head of the Physics Department

1960 - 1968:
James Blair High School, Willamsburg, Virginia
Guidance Counselor, Physics and Mathematics Teacher.

1957 - 1960:
Institute of Textile Technology, Charlottesville, Virginia
Research Associate in Statistics Department and Thesis Advisor to Graduate Students in Textile Technology

ABSTRACT

Most institutions of higher education are interested in enrollment projections because they are closely related to institutional goals and missions and, are, therefore, essential to financial and program planning at every level. This study was undertaken to determine if relevant factors could be identified and used in a statistical forecasting model to project enrollments in a multidimensional urban community college within the accuracy limitations imposed by a state such as Virginia (who requires State institutions of higher education) to project their enrollment within ± 1 percent.

Two general types of statistical forecasting models, causal and extrapolation models were explored for use in forecasting fall and summer headcount, and total FTE enrollments within the prescribed accuracy limits. The relevant factors for possible inclusion in the models were identified from previous studies and a student flow model for the institution. The relevant factors used in the final models were selected on the basis of simple correlation coefficients, the mean square error, and average error as variables were added and removed from the models.

The optimum fall and summer headcount forecasts were produced by a combination time-series and multiple regression model. The independent variables used in fall and summer headcount forecasts were a seasonal factor, a time-trend factor, and national economic indicators. In the optimum total FTE forecast, produced by a multiple regression model, the relevant factors were full-time enrollment shifted forward three years and national economic indicator shifted forward three years. The basis for acceptance or rejection of the models was made in context with the fiscal system of the Commonwealth of Virginia for the distribution of public funds to the state colleges and universities. The fiscal system was established primarily to provide a basis for financial planning. Forecasting models were produced for 1 year for fall headcount enrollment and for 2 years for summer headcount and total FTE enrollment within ± 1 percent.

On the basis of this study certain general conclusions were reached: the large variations between national enrollment projections resulted from different assumptions; enrollment projections have been too generalized for institutions with diverse goals and objectives; present data bases are inadequate to produce accurate enrollment projections; and most projections are not sufficiently reliable for planning purposes. More specific conclusions reached were: state data bases are inadequate for multidimensional institutions; removing quarterly seasonal variations permits the identification of relevant factors; traditional projection models such as the cohort survival and Markov are not applicable in multidimensional institutions such as community colleges; models such as time-series and multiple regression can be developed to accurately project enrollments for less than two years; the current limits of accuracy for Virginia multidimensional institutions are unrealistic; verification of the accuracy of prediction models is valuable for evaluating forecasting models; and models for multidimensional institutions must be revised periodically because relevant factors are constantly in flux.