DOCUMENT RESUME

ED 431 363 HE 032 128

AUTHOR	Adelman, Clifford
TITLE	Answers in the Tool Box. Academic Intensity, Attendance Patterns, and Bachelor's Degree Attainment.
INSTITUTION	National Inst. on Postsecondary Education, Libraries, and Lifelong Learning (ED/OERI), Washington, DC.
REPORT NO	PLLI-1999-8021
PUB DATE	1999-06-00
NOTE	141p.
AVAILABLE FROM	ED Pubs, P.O. Box 1398, Jessup, MD 20794-1398; Tel: 877-433-7827 (Toll Free).
PUB TYPE	Numerical/Quantitative Data (110) -- Reports - Research (143)
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	Academic Achievement; *Academic Persistence; *Bachelors Degrees; College Admission; College Outcomes Assessment; High Schools; Higher Education; Longitudinal Studies; *Predictor Variables; Remedial Instruction; Secondary School Curriculum; Statistical Surveys; Student Surveys; Tables (Data); Time To Degree
IDENTIFIERS	*Graduation Rates; *High School and Beyond (NCES)

ABSTRACT

This study of degree completion followed a national cohort of students from 10th grade (in 1980) through 1993 (High School and Beyond/Sophomore data base). Data included high school and college transcripts, test scores, and surveys. Using linear regression techniques, the study identified a model that accounts for about 43 percent of the variance in degree completion. The two most important variables identified were, first, academic resources a composite measure of the academic content and performance the student brings from secondary school, and, second, continuous enrollment. Major conclusions include the following: (1) since many students attend multiple institutions, institutional graduation rates are not very meaningful; (2) college admissions formulas that emphasize test scores and high school grade point average (rather than academic intensity and curriculum quality) are likely to produce lower degree completion rates; and (3) type and amount of remediation matters in relation to degree completion. Part 1 of the report constructs an index of student academic resources; Part 2 analyzes and/or reconstructs the major variables; Part 3 examines new attendance patterns and their significance; and Part 4 builds statistical models to explain findings. Five appendices include technical notes, additional tables, and examples. (Contains 125 references.) (DB)

Answers in the Tool Box

Academic Intensity, Attendance Patterns, and Bachelor's Degree Attainment

U.S. Department of Education
Office of Educational Research and Improvement

Answers in the Tool Box:

Academic Intensity, Attendance Patterns, and

Bachelor's Degree Attainment

Clifford Adelman
Senior Research Analyst
Office of Educational Research and Improvement
U.S. Department of Education

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
C. Kent McGuire
Assistant Secretary

National Institute on Postsecondary Education, Libraries, and Lifelong Learning
Carole Lacampagne
Director

Media and Information Services
Cynthia Hearn Dorfman
Director

June 1999

To obtain copies of this report (while supplies last) or ordering information on other U.S. Department of Education products, call toll-free 1–877–4ED–Pubs (877–433–7827) or write

Education Publications Center (ED Pubs)
U.S. Department of Education
P.O. Box 1398
Jessup, MD 20794–1398

TTY/TTD 1–877–576–7734
FAX 301–470–1244

Contents

Acknowledgments

Whatever value one finds in this monograph owes much to the insistent and insightful world-class reviewers of its drafts: Dennis Carroll, Paula Knepper, Drew Malizio, Ann Mullen, and Jeffrey Owings of the National Center for Education Statistics, Karl Alexander of the Johns Hopkins University, Alberto Cabrera of the Pennsylvania State University, Aaron Pallas of Michigan State University, and Vincent Tinto of Syracuse University. They all gave generously of their time and care, and were delightfully uncompromising in their requirements for improvement. One is blessed by such teachers. In the final analysis, they are collaborators.

Earlier versions of Parts I and III of this monograph were presented as research papers at the 1998 Forum of the Association for Institutional Research (Part I) and the 1998 Annual Conference of the Association for the Study of Higher Education (Part III). On these occasions, colleagues from both organizations poked, prodded, and encouraged the completion of the more embracing story-line contained in these pages. I am indebted to them as I am to the challenges and proddings of Lawrence Gladieux and Scott Swail of the Washington office of the College Board, where I was privileged to serve as a visiting fellow in 1998.

Thanks, too, to my colleagues Jim Fox, Harold Himmelfarb, and Joe Teresa for their readings of a very rocky first draft, and for questions that told me where second thoughts and greater clarity were necessary.

Smaller pieces of the analysis, along with some of the "tool box" recommendations, have been aired previously in the general and trade press, and I thank *Change* magazine, *University Business,* the *Chronicle of Higher Education*, and the *Washington Post* for bringing these messages to broader publics than this more technical version will reach.

Executive Summary

Answers in the Tool Box is a study about what contributes most to long-term bachelor's degree completion of students who attend 4-year colleges (even if they also attend other types of institutions).

Degree completion is the true bottom line for college administrators, state legislators, parents, and most importantly, students—not retention to the second year, not persistence without a degree, but completion.

This study tells a story built from the high school and college transcript records, test scores, and surveys of a national cohort from the time they were in the 10th grade in 1980 until roughly age 30 in 1993. The story gives them 11 years to enter higher education, attend a 4-year college, and complete a bachelor's degree. In these respects–based in transcripts and using a long-term bachelor's degree attainment marker--this story is, surprisingly, new.

This study was motivated by four developments in higher education during the 1990s:

(1) The growing public use of institutional graduation rates as a measure of accountability, and the tendency in public policy and opinion to blame colleges for students' failure to complete degrees and/or for failure to complete degrees in a timely manner.

(2) An ever expanding proportion of high school graduating classes entering postsecondary education, and new federal policies encouraging even more students to enter or return to higher education. Our system is being challenged simply to maintain, let alone improve, college graduation rates.

(3) The increasing tendency, overlooked in both policy and research, for students to attend two, three, or more colleges (sometimes in alternating patterns, sometimes simultaneously) in the course of their undergraduate careers.

(4) The rising heat of disputes involving admissions formulas at selective colleges where affirmative action policies have been challenged. These disputes, carried into the media and hence dominating public understanding, involve two indicators of pre-college attainment—grades/class rank *versus* test scores—without any reference to high school curriculum and its role in the degree completion rates of the mass of minority students.

The story of what contributes most to bachelor's degree attainment works toward six ordinary least squares regression equations that progressively add blocks of key variables following the progress of students from high school into higher education and through the first true year of attendance. The penultimate model (the fifth in the series) accounts for about 43 percent of the variance in bachelor's degree completion [p. 74]. The sixth equation simply indicates that one hits a plateau of explanation at this point. For a story-line such as this, 43 percent is a very high number. A five-step logistic regression then provides both a dramatic underscoring of the principal findings and some enlightening variations.

There are 11 variables in the penultimate linear regression model. The two most important variables, accounting for the bulk of the model's explanatory power are:

○ "Academic Resources," a composite measure of the academic content and performance the student brings forward from secondary school into higher education. This measure is dominated by *the intensity and quality of secondary school curriculum* [Part I and Appendix C].

○ Continuous enrollment once a true start has been made in higher education.

In the logistic version of the penultimate model, the same 11 variables (out of 24) are statistically significant, but those displaying the strongest relationships to degree completion (the highest "odds ratios") are all post-matriculation phenomena: continuous enrollment, community college to 4-year college transfer, and the trend in one's college grades.

Among the 11 variables, the following are not usually found in similar analyses:

○ Proportion of undergraduate grades indicating courses the student dropped, withdrew, left incomplete, or repeated. [pp. 54-56]
○ A final undergraduate grade point average that is higher than that of the first "true" year of attendance. [pp. 72-73]
○ Parenthood prior to age 22. [pp. 37-38]
○ Whether the student attended more than one institution and did *not* return to the first institution of attendance, a situation that includes, but transcends, the classical community college to 4-year college transfer pattern. [p. 46]

The only demographic variable that remains in the equation at its penultimate iteration is socioeconomic status, and by the time students have passed through their first year of college, SES provides but a very modest contribution to eventual degree completion. No matter now many times (and in different formulations) we try to introduce race as a variable, it does not meet the most generous of threshold criteria for statistical significance.

Selected Findings

High School Background

○ High school curriculum reflects 41 percent of the academic resources students bring to higher education; test scores, 30 percent; and class rank/academic GPA, 29 percent [p. 21]. No matter how one divides the universe of students, the curriculum measure produces a higher percent earning bachelor's degrees than either of the other measures [p. 15]. The correlation of curriculum with bachelor's degree attainment is also higher (.54) than test scores (.48) or class rank/GPA (.44) [p. 19].

• The impact of a high school curriculum of high academic intensity and quality on degree completion is far more pronounced—and positively—for African-American and Latino students than any other pre-college indicator of academic resources. The impact for African-American and Latino students is also much greater than it is for white students [pp. 84-86]

• Of all pre-college curricula, the highest level of mathematics one studies in secondary school has the strongest continuing influence on bachelor's degree completion. Finishing a course beyond the level of Algebra 2 (for example, trigonometry or pre-calculus) more than doubles the odds that a student who enters postsecondary education will complete a bachelor's degree. [pp. 16-18]

• Academic Resources (the composite of high school curriculum, test scores, and class rank) produces a much steeper curve toward bachelor's degree completion than does socioeconomic status. Students from the *lowest* two SES quintiles who are also in the *highest* Academic Resources quintile earn bachelor's degrees at a higher rate than a majority of students from the top SES quintile. [pp. 24-25]

• Advanced Placement course taking is more strongly correlated with bachelor's degree completion than it is with college access. [pp. 19-20]

• Graduating from high school "late" does not influence bachelor's degree completion provided that one enrolls in higher education directly following receipt of the diploma and attends a 4-year college at some time [p. 68-69].

College Attendance Patterns

• The proportion of undergraduate students attending more than one institution swelled from 40 percent to 54 percent (and among bachelor's degree recipients, from 49 to 58 percent) during the 1970s and 1980s, with even more dramatic increases in the proportion of students attending more than two institutions. Early data from the 1990s suggest that we will easily surpass a 60 percent multi-institutional attendance rate by the year 2000. [pp. 42-45]

• Students beginning in highly selective 4-year colleges and those starting out in open door institutions have the highest rates of multi-institutional attendance, though for very different reasons. [p. 45]

• The number of institutions attended by students has no effect on degree completion. [p. 68].

• The fewer schools attended, the more likely the student was enrolled continuously, and the less likely a 4-year college was part of the attendance pattern. [p. 48] Yet 70 percent of the students who attended a 4-year college at any time were continuously enrolled. [p. 54]

• Sixteen (16) percent of postsecondary students (and 18 percent of bachelor's degree completers) engaged in alternating or simultaneous enrollment patterns. Some 70 percent of this group attended three or more institutions. [pp. 45-46]

• Some 40 percent of students who attended more than one institution crossed state lines in the process, and their bachelor's degree completion rate was higher than that for multi-institutional students who remained within state borders. [p. 49-50]

• Students who expected to earn a bachelor's degree, started in a 2-year institution, but never attended a 4-year college have a lower SES profile and a considerably lower academic resources profile than students with the same expectations and starting point but who did attend a 4-year school. Family income, however, plays no role in the different attendance patterns of these students. [pp. 57-59]

Degree Completion

• For students who attend 4-year colleges at some time, the only form of financial aid that bears a positive relationship to degree completion after a student's first year of college attendance is employment (principally College Work-Study and campus-related) undertaken (a) while the student is enrolled and (b) for purposes of covering the costs of education. [pp. 64-65]

• The long-term national *system* bachelor's degree completion rate by age 30 for all *students* who attend 4-year colleges is 63 percent; for all those who earn more than 30 credits, the rate exceeds 70 percent. [pp. 28-29] For those who start in highly selective colleges, the rate exceeds 90 percent. [p. 52]

• While only 26 percent of students who began their undergraduate careers in community colleges formally transferred to 4-year institutions, their bachelor's degree completion rate was over 70 percent. [pp. 53-54] The classic form of transfer, in which the student earns at least a semester's worth of credits before moving to the 4-year college, produces a very high likelihood of bachelor's degree completion. [pp. 80-82]

• The mean elapsed time to complete a bachelor's degree for this cohort was 4.72 *calendar* years, or 5 full academic years. For students in the highest quintile of pre-college academic resources, the mean time was 4.45 calendar years. For students who were continuously enrolled, it was 4.33 calendar years. [Appendix D]

- Thirty-nine percent of 4-year college students who were assigned to remedial reading courses completed bachelor's degrees, compared with 60 percent of students who took only one or two *other* types of remedial courses, and 69 percent of those who were not subject to remediation at all. [p. 74]

- Students who attend 4-year colleges and who earn fewer than 20 credits in their first calendar year of postsecondary experience severely damage their chances of completing a bachelor's degree. [pp. 70-71, 81]

Conclusions That Follow from These Findings:

- When nearly 60 percent of undergraduates attend more than one institution and 40 percent of this group do not complete degrees, institutional graduation rates are not very meaningful. It is not wise to blame a college with superficially low graduation rates for the behavior of students who swirl through the system.

- Analysis of institutional effects on degree completion is compromised when students attend two or more institutions. One wastes precious research time trying to figure out which type of experience in institution X had an impact if the student also attends institutions Y and Z. There are some exceptions to these principles, e.g. when the second institution involves a study abroad semester.

- When the academic intensity and quality of one's high school curriculum is such a dominant determinant of degree completion, and both test scores and (especially) high school grade point average or class rank are so much weaker contributors to attainment, college admissions formulas that emphasize test scores and (especially) high school grade point average or class rank are likely to result in lower degree completion rates.

- The type and amount of remediation matters in relation to degree completion. Increasingly, state and local policy seeks to constrict—if not eliminate—the amount of remedial work that takes place in 4-year colleges. But there is a class of students whose deficiencies in preparation are minor and can be remediated quickly without excessive damage to degree completion rates.

What We Learned: Variables to Discard

Examples of stock building-block variables that are discarded because of weak architecture:

- Highest level of parents' education. As reported by students, these data are uneven and unreliable. In the most recent of the national longitudinal studies, the *highest* degree of agreement between students and parents on this score was 72 percent in the case of fathers with "some college." One out of six students would not even venture a guess as to their parents' education. [pp. 37-38]

• "Persistence" defined in temporal terms, e.g. from the 1st to 2nd year of college. Transcripts reveal an enormous range in the *quality* of arrival at the putative 2nd year: some 30 percent of those who were "retained" or "persisted" arrived with either less than 20 credits or 3 or more remedial courses. [p. 27]

• "Academic track" (sometimes called "college preparatory") in secondary school curriculum, whether reported by students or by schools. When the transcripts for a third of the students on the "academic track" show 8 or fewer Carnegie units in core academic subjects, it is obvious that the transcripts—not the label—must be the source of judgment. [p. 10]

• "Part-time" enrollment in postsecondary education. Students change status from term to term. Part-year enrollment may be more important than light credit loads. Most importantly, students change status within a given term, by dropping, withdrawing from, or leaving incomplete large portions of their credit loads. The "DWI Index" (ratio of drops/withdrawals/incompletes to total courses attempted) derived from transcript records is far more important than what the student says in an interview about full-time/part-time status. [pp. 54-56]

. . .and Variables Reconstructed

• Academic intensity and quality of high school curriculum. This is the most elaborate construction in the study. It includes Carnegie units in 6 academic areas, accounts for highest mathematics studied, remedial work in English and math, and advanced placement. The construction results in a criterion-referenced scale with 40 gradations. [pp. 12-14, and Appendix C]

• Educational aspirations. Traditionally defined on the basis of a single question asked in the senior year of high school. Reconstructed on the bases of 6 pairs of questions asked in both 10th and 12th grades, and on the principles of consistency and level. The result is a statement of "anticipations," not "aspirations." [pp. 33-36]

• First institution/date of attendance in postsecondary education. Redefined from college transcript data to exclude false starts and incidental attendance in the summer following high school graduation. [pp. 44-46]

• Transfer. The classic form of community college to 4-year college transfer is now a sub-set of a larger multi-institutional attendance pattern universe defined here in terms of 9 sets of institutional-type combinations. Transfer as we knew it has been replaced by what one might call "portfolio building." [pp. 46-49] But the classic form of transfer is an extremely effective route to bachelor's degree completion.

What We Learned: Principles to Guide Research and Evaluation

- Institutions may "retain" students, but it's *students* who complete degrees, no matter how many institutions they attend. So follow the student, not the institution.

- Common sense can tell us what's likely to be important at every step toward the degree. A fierce empiricism will validate common sense.

- Before one accepts a variable simply because it has been used for decades or because a federal agency paid for it, one must examine the bricks and mortar of that variable very carefully. Where the architecture is faulty, the data must be fixed or the variable discarded—or one will never tell a true story.

- We should not compute bachelor's degree attainment rates for people who never set foot in a bachelor's degree-granting institution.

- The most useful data lie in the details, not the generalities.

The monograph concludes with "tool box" recommendations to those who execute policy regarding both pre-college opportunity-to-learn and post-matriculation advisement. The tool box metaphor is a logical consequence of the analysis. It says that if we are disappointed with uneven or inequitable outcomes of postsecondary education, we must focus our efforts on aspects of student experience that are realistically subject to intervention and change.
We do not have tools to change intentions or perceptions, or to orchestrate affective influences on students' decisions. The events of students' life course histories through their 20s lie largely beyond the micromanagement of collegiate institutions. But we do have the tools to provide increased academic intensity and quality of pre-college curricula, to assure continuous enrollment, to advise for productive first-year college performance, and to keep community college transfer students from jumping ship to the 4-year institution too early.

The recommendations thus address dual enrollment, direct provision of secondary school curriculum by college instructors, an 11-month rolling admissions cycle for all 4-year colleges, using Internet situated courses to keep college students continuously enrolled (even for one course), implementation of institutional policies restricting the extent of course withdrawals/incompletes/repeats, realistic credit loads, and advisement that is both sensitive and sensible.

The story and its analyses are derived from and apply to a cohort whose history covers the period 1980-1993. There is another and more contemporary cohort whose history, beginning in 1988, is still in progress. Will the story-line change? Will the analyses be validated? Will we have attained greater equity in degree-completion rates for minority students? Have attendance patterns become even more complex, and more oriented toward competences and certifications as opposed to degrees? Only a full data-gathering for this cohort in the year 2000 and the collection of its college transcripts in 2001 will tell.

On Reading Tables in This Study

Most of the tables in this study include a "universe statement" that tells the reader exactly who is included in the statistics, along with the weighted number of students in that universe.

The "universe statement" can be found in the notes for the tables at issue. It is a most important guide for interpreting a table.

For example, many tables include an estimate for bachelor's degree completion rates. No two of these estimates are the same, and the reasons lie in the definition of the universe.

The definition serves as more than a documentation and guide: it is an academic courtesy. It allows other researchers who wish to investigate different hypotheses or different configurations and definitions of variables to replicate the universe and thus report more accurately how their conclusions either reinforce or differ from those stated here.

Introduction: Departing from Standard Accounts of Attainment

This study analyzes the relationship between the academic resources students bring to college, their long-term attendance patterns, and their degree completion rates.

The study is both related to and departs from a number of honored lines of research on the determinants of educational attainment. It diverges from previous research on attainment principally by emphasizing the details of students' high school and college curricula and academic histories that are available from transcript records. Its principal data are drawn from the new (1998) restricted edition of the High School & Beyond/Sophomore cohort files (hereafter referred to as the HS&B/So)[1]. This longitudinal study followed a national sample of students from the time they were in the 10th grade in 1980 to roughly age 30 in 1993[2].

In round numbers, of the high school graduates in this cohort, 65 percent attended some form of postsecondary education and 40 percent attended a 4-year college by age 30. These are basic "access rates." Of the group attending 4-year colleges at some time, 63 percent earned a bachelor's degree.

While the 63 percent completion rates sounds impressive for a mass system of higher education, it masks an unhappy differential in degree completion rates by race/ethnicity. Furthermore, we have since reached a 75 percent access rate (Berkner and Chavez, 1997), and, in the late 1990s, our national policies have invited an even higher percentage of high school graduates into postsecondary education. Simply to maintain—let alone improve—our long-term degree completion rate will take a great deal of effort. We need guidance.

So this study asks a simple question:

> What contributes most to bachelor's degree completion of students who attend 4-year colleges at any time in their undergraduate careers?

The answers to that question help us develop strategies to address anomalies, paradoxes, and disappointments in educational attainment after high school. The answers suggest what tools to put in our tool boxes and where to take them. The answers advise us how to use the tools in an environment of changing student enrollment behavior. The answers clearly instruct us as to what is important in research on this issue and what is no longer so important. The answers may not be exhaustive, but without them, there is no tool box.

The HS&B/So is the second of three national age-cohort longitudinal studies carried out under the design and sponsorship of the National Center for Education Statistics (NCES). From time to time in this monograph, the other two studies will be invoked: the National Longitudinal Study of the High School Class of 1972 (hereafter referred to as the NLS-72) and the National Education Longitudinal Study of 1988 (hereafter referred to as the NELS-88). For the reader's convenience in bench marking, think of these three studies as following the scheduled high school graduating classes of 1972, 1982, and 1992. The last of these studies, the NELS-88, is still in progress. Occasionally, too, data from NCES's Beginning Postsecondary Students longitudinal study of 1989-1994[3] will enter the discussion.

Why Are We Asking Such an Obvious Question?

While employers increasingly use "college degree" as a screening device in hiring (National Center on the Educational Quality of the Workforce, 1998), and legislatures everywhere ask for evidence of "graduation rates," the research literature devotes infinitely more attention to access than to degree completion. How strange! And stranger, still when, in spite of all we know about recurrent education and delayed entry to higher education, in spite of all the public displays of 30-year-olds returning to complete degrees they abandoned to have children or start businesses, how little research uses *long-term* degree completion as its time-frame. In the country of the second and third chance, our legislation and our research ask us to hurry up and get it over with, and judge both individuals and institutions negatively if they fail to get it over with fast.

Yet, as we will see, the winds have changed, and both our legislation and our research have yet to acknowledge that change: going to college in the 1990s means something very different from what it meant 20 years ago. Unless we recognize these changes, the higher education enterprise will drift like a ship in the horse latitudes. One reason for asking the simple question is to help policy re-navigate to find the winds and the new currents of student attendance patterns.

Yet another reason for asking the basic question lies in contemporary policy disputes involving admissions formulas at selective institutions, principally as a by-product of a new dispensation for affirmative action. The heat of those disputes has unfortunately been raised by a dubious argument over two indicators of pre-college attainment—grades/class rank *versus* test scores—that make no reference whatsoever to curricular content. We owe it to students, and to minority students in particular, to assess the most profitable paths to degree completion in *any* institution. This obligation decrees that we explore the potential power of secondary school curriculum to set students on trajectories that will culminate in a satisfactory ending for them and for the society writ large. We talk a great deal in policy about school-college connections and collaborations. Test scores and class rank have little to do with those connections and collaborations. Curriculum has everything to do with them.

In these pages, then, the reader will see a great deal of college attendance patterns and high school curriculum.

Structure of This Monograph

Parts I-III of this monograph set up the principal touchstones for the story that seeks to answer our basic, simple question. We begin by demonstrating how to construct an index of student academic resources (Part I). The analysis then sorts through a few major variables frequently found in analyses of persistence and degree completion, discarding some as unreliable, and reconstructing others (Part II). Finally, we open up the new world of attendance patterns in higher education, and explore its significance for both research and policy (Part III). College administrators, state policy-makers, researchers, and journalists often ask "How did you get that number?" or "What do you mean by that variable?" Watching the construction and testing of a variable provides the answers. The major variables in this analysis are not buried in appendices, though to make life easier for the general reader, some of the technical material and comments on the literature have been placed in endnotes.

Part IV builds two series of statistical models to explain what made a difference in bachelor's degree completion by age 30 for the students in the HS&B/So cohort. First, a series of linear Ordinary Least Squares regressions seeks to explain how much of the variance in bachelor's degree attainment can be attributed to different background characteristics, achievement, and experiences when all other variables in an equation are held constant. Each model in the series takes up a successive stage in students' life histories, dropping those variables that don't make a difference and adding new sets of variables until we reach a plateau of explanation. Second, the same models are tested using logistic regressions in a manner suggested by Cabrera (1994) to provide a different type of portrait of the results. The concluding section of the monograph explores the contemporary significance of the findings in terms of achieving a greater degree of equity in degree completion rates.

"Academic Resources" and the Dow Jones

I did not invent the term, "academic resources," as used in this paper. Credit goes principally to Karl Alexander and his various associates over nearly two decades of research on the paths from secondary to postsecondary education (e.g. Alexander and Eckland, 1973; Alexander and Eckland, 1977; Thomas, Alexander, and Eckland, 1979; Alexander and Cook, 1979; Alexander, McPartland and Cook, 1981; Alexander and Cook, 1982; Alexander, Riordan, Fennessey and Pallas, 1982; Alexander and Pallas, 1984; Alexander, Pallas, and Holupka, 1987; Alexander, Holupka, and Pallas, 1987; Pallas and Alexander, 1983), and this material will be cited frequently. Alexander and his colleagues persistently demonstrated that the power of a student's academic background overwhelms the predictive power of demographic variables (gender, race, socioeconomic status) in relation to test performance (Alexander and Pallas, 1984), college attendance (Thomas, Alexander, and Eckland, 1979) and, in one study, college completion (Alexander, Riordan, Fennessey and Pallas, 1982), yet few higher education researchers pay much attention to this body of literature. At the same time, what Alexander and his colleagues mean by "student academic background" calls out for revisiting and reconstruction, and one of the purposes of this study is to expand, deepen, and test the concept of student academic resources in light of both transcript data and long-term paths *through* postsecondary education to degree completion.

Indeed, while most related research focuses on access or year-to-year retention, the dependent variable in this study is completion of bachelor's degrees, the Dow Jones Industrial Average of U.S. higher education. The reasons for focusing on degree completion relate principally to equity issues in an age when 65 percent of high school graduates enter higher education directly from high school and 75 percent enter within two years of high school graduation (Berkner and Chavez, 1997). While the "college access gap" between whites and blacks and whites and Latinos has closed from the 11-15 percent range to 5 percent over the past two decades, the degree completion gap remains stubbornly wide at 20 percent or higher (Smith *et al*, 1996, p. 25), and it behooves us to inquire into this unhappy paradox in somewhat different directions than have been followed in the past.

The reasons for focusing on attendance and curriculum patterns in the analysis are that we have become as mobile and consumeristic in higher education as we are in the rest of our lives, that we no longer stay in one place for prescribed periods of time, and that we feel free to mix various life activities in whatever order we wish—or whatever order is made necessary by other life commitments and circumstances (Blumberg, Lavin, Lerer, and Kovath, 1997). What one studies may thus be more important than the many places at which one studies it.

Our abiding interest in research on retention and completion is to discover those aspects of student and institutional behavior that actually can be changed to improve the odds of attainment, even though our definition of attainment may be different from that of some students (Tierney, 1992). We look for concrete and practical suggestions that can be assigned to particular individuals and groups to carry out, not the generalized, abstract flourishes that Orwell (1949) called "soft snow," and that we witness at the conclusion of too many research articles that expend their energy on complex statistical modeling.

Because we seek behavior that can be changed, our research must focus on conditions that are realistically subject to manipulation by people in the best positions to do so, people who can use the tool box. For example, some research has demonstrated the strong role of parents, peers, and significant others in student decisions to attend college, choose a particular college, choose a particular major, and choose to persist (Bean, 1982). More recent research has demonstrated that reputation and location are criteria that overwhelm all others (influence of parents, peers, etc. included) in choosing a postsecondary institution (Choy and Ottinger, 1998). While high school counselors and teachers can work with parents in matters of preparing students for college and encouraging application, there is very little anybody else can do to orchestrate these external players in terms of affective influences on *post-matriculation* student enrollment and persistence behaviors. There is even less that one can do within the expanding patterns of transfer and multi-institutional attendance that the HS&B/So 11-year history (to age 30) reveals and that will be detailed below. For those beyond the age of 30, the decision to return to complete degrees begun earlier is even more influenced by complex interactions of external and personal factors (Smart and Pascarella, 1987; Spanard, 1990). Events in life-course history such as changes from dependent to independent status, marriage and divorce, and increases in the number of children in a household lie beyond the micro-management of higher education faculty and staff[4].

The Tenor of History

There is a tenor to the approach and methodology of this study that also should be posited at the outset because it departs from reigning models. The tenor is that of exploratory historical investigation, and thus inevitably conditions what I regard as credible evidence, what meets the criteria for statistical relationships, what type of regression analysis is best suited to chronological story-telling, and what we might call "the problem of the typical."

History is a fiercely empirical discipline. The evidence it assembles is all tangible: artifacts, diaries, parish records, letters, communiques, e-mails, texts, photographs, recordings, ruins, taped interviews, dictionaries, maps, ship's manifests and logs, etc. Truth often lies in—and can be extracted from—the details. Historians do not design or conduct surveys (their subjects are often dead, so surveys are a moot methodology). Rather, they will find surveys and treat them as texts (see, for example, Clubb, Austin, and Kirk, 1989). They are interested, foremost, in the traces of human behavior, "the marks, perceptible to the senses, which some phenomenon . . . has left behind" (Connerton, 1989, p. 13). Thus, unobtrusive evidence is of paramount value in history. While historians may speculate about the meaning and significance of that evidence, they treat it as authoritative, even when they take samples of the evidence as representing characteristics of populations (Haskins and Jeffrey, 1990. chapter 4). They may discover that the evidence was contrived, but they then will treat the fact of contrivance as equally authoritative.

What does this fierce empiricism mean for interpreting a data set that was prepared for the National Center for Education Statistics (or any other federal agency, for that matter)? Simply because we paid someone for gathering and coding the data does not mean the data were handed down from Mount Sinai and must never be questioned. The data set consists of historical evidence, which is "in no sense privileged" (Connerton, 1989, *loc cit*). Every case of every variable requires examination. Anomalies are subject to multiple examinations. Editorial adjustments and corrections are made only under the strictest decision rules. But these adjustments and corrections *must* be made or we will never tell a true story. The practice is called historical reconstruction.

While accounts of these editorial processes and their decision rules have been published elsewhere (Adelman, 1997; Adelman, 1995), an example would be helpful. Assume a student for whom we have college transcripts and a full post-secondary history beginning in the fall of 1982. But the secondary school record for this student appears strange and spotty. We do not know what the coders were looking at when they entered data from this student's high school transcript back in 1982, but there appear to be only 6 Carnegie units on the transcript, no indication that the student changed high schools (which might explain a truncated record), and no indication the student ever studied mathematics or foreign language. The college transcript records, however, show that in the fall semester of 1982, the student entered a flagship campus of a state university, earned a B+ in calculus 3 and a B in Russian conversation and composition 5. On the basis of this information alone[5] we can reasonably revise our record of the student's high school transcript to include 3 units of mathematics through pre-calculus and

2 units of foreign language. Using independent sources, we know the requirements for high school graduation in the student's home state (Medrich, Brown and Henke, 1992). The flagship campus of the state university would not accept the student unless he/she was a high school graduate, hence had met those requirements. By the time we are done, our record of the student's high school transcript is a lot more accurate than what we originally received. The secondary school records of approximately 18 percent of the HS&B/So students were subject to this type of adjustment.

Editing in this manner may involve inference, but not what statisticians call imputation. It does not assign specific behaviors or attainments to masses of people on the basis of the intrinsic characteristics of those people. It does not say, "because you look like all other people of a certain configuration of characteristics and because your survey form is missing transcripts, we are going to assign you the degrees, majors, test scores, etc. of those other people." At no point in editorial work on a data set will a historian make such assumptions and impute characteristics to individuals on the basis of group models.

Explanation More Than Prediction

The second departure from reigning modes of analysis of postsecondary careers derives from one of the most fundamental lessons of history: while stories may repeat themselves, they never do so in the same way. Even when they employ quantitative methods, historians are not in the prediction business, and, with rare exceptions, do not worry about directional causality[6]. Researchers have spent the past two decades attempting to squeeze every drop of predictive blood from the data on college access and persistence. They have consumed thousands of journal pages with arguments over the comparative power of different statistical models: factor analysis, structural equations such as LISREL, logistic regression, weighted least squares regression, probit, etc. (Dey and Astin, 1993). By the time we are done reading this library shelf, contrary predictions often arise, and the point is lost on anyone who might use the information. While employing statistical models commonly used in prediction, this study is less interested in forecasting the behavior of future students than in explaining what made a difference for past students.

To be sure, the story may provide guidelines for thinking about the experience of future cohorts, but the groups will inevitably differ. A statistical model derived from a class that entered higher education in 1968, when the majority of students were middle-class white males who enrolled full-time and directly from high school, may reveal relationships that are worth exploring with contemporary populations, but is still unique to its time and circumstance. The proper form of a sentence stating the conclusion of an equation for such a cohort might be, for example, "the socioeconomic composition of one's high school class had a greater net impact on attainment for this group than the selectivity of the first college attended." That sentence is, in fact, a re-write of a major conclusion reached by Alexander and Eckland (1977).

The story we tell about a cohort rests on the assumption that what we observe is representative, or "typical," of that population. One of the principal reasons for performing statistical tests, in fact, is to demonstrate that the story line and its components did not come about by chance, and that there is far more coherence than chaos. The task is analogous to that faced by historians attempting to determine what is typical of a particular culture or sub-culture during a particular period. The notion of "typical" may involve a range or array of behaviors, attitudes, conditions—and these are derived from the traces of artifacts, records, and the style of texts.

Searching through the details of these remains, one cannot determine what is "typical" by collapsing variables into categories at such a level of aggregation that a constructive story-line is impossible to detect. For example, if one is going to describe the geographic "region" of potential college students (St. John, 1991; St. John and Noell, 1989) and under the conviction that the quality of student preparation is determined by geographic region, then four "regions" consisting of 13-14 states is much too large an aggregation. There are nine (9) Census divisions one can invoke, and the more promising analytic combination is that of Census division by urbanicity of high school (urban, suburban, rural), yielding 27 cells. If one wants to know the geographic origins of students taking more than two college courses at remedial levels so that one can address the comparative severity and distribution of the problem of remediation, 27 combinations offer some compelling suggestions[7]. Four regions do not help us take our tool boxes to the places they are needed. As Hearn (1988) noted, "it is in the details that the most precise, and most useful, answers lie" (p. 173).

What Evidence Do We Use? The Case of Student Self-Reports

How do we know that students were taking remedial[8] courses in college, let alone what kind of remedial courses? Do we ask the students, and, if so, how (are you enrolled this term in a remedial course? were you ever enrolled . . .?)? Do we use a cross-sectional survey of registrars (Lewis and Farris, 1997)? Or do we use college transcripts, and trace remedial problems back through high school transcripts? Let us briefly compare what we find from each of these methods.

- Only 15 percent of the *students* in the five-year longitudinal study, Beginning Postsecondary Students, 1989-1994, told us that they took at least one remedial course during the period they were in postsecondary education. Source: Data Analysis System (DAS), BPS90

- *Institutional officials* told the National Center for Education Statistics' Postsecondary Education Quick Information Service that 29 percent of *freshmen* were taking remedial courses in the fall term of 1995 (Lewis & Farris, 1997).

- The postsecondary *transcripts* of the HS&B/So cohort show that 46 percent of all students took at least one remedial course during their undergraduate years (1982-1993).

Granted, these are three different surveys with different time periods. But the discrepancies between the unobtrusive evidence (transcripts), second-party accounts, and student testimony are simply too great for comfort[9], and it is worth further demonstration of this problem.

Two tables should drive home the virtues of unobtrusive evidence. Table 1 demonstrates the disparities between students' claims to degree attainment and the evidence of their transcript records. What do we see in this table? (1) about 7 percent of those who claim to earn a bachelor's degree or higher have earned, at best, an associate's degree; (2) some people do not understand the question about highest degree, and claim *less* than the evidence shows they have earned; (3) the concept of a "certificate" is very slippery, and people will try to claim at least some minimum postsecondary credential as psychological compensation for their time; and (4) because there was a 12-15 month gap between the date of the 1992 survey interview and the period of 1993 transcript receipt, it appears that some students in graduate school expressed *expectations* for degree completion in 1992 that were not realized by 1993.

Table 1–Discrepancies between highest degree claimed and highest degree earned by students in the High School & Beyond/Sophomore cohort

	Highest Degree Earned by 1993 (Transcript Evidence)					
	None	Certif-	Assoc	Bachel	Gradu	% of All
Highest Degree Claimed in 1992:						
None	93.0%	2.2%	1.4%	2.5%	0.4%*	36.8%
Certificate	48.9	49.2	1.0	0.7*	0.2*	14.3
Associate's	16.3	18.2	63.0	2.4*	0.1*	12.5
Bachelor's	4.7	0.8*	1.6	75.0	18.0	30.6
Graduate	2.7*	---	0.9*	9.4	87.0	5.8
% of Earners	45.0	10.4	9.1	24.8	10.7	100.0

Notes: (1) Rows add to 100.0%; (2) Universe consists of all high school graduates in the HS&B/So who answered the 1992 survey question concerning highest earned degree, and for whom transcript evidence was available; (3) Weighted N=2.29M; (4) *=Low N cells. **SOURCE**: National Center for Education Statistics: High School & Beyond/Sophomore cohort, NCES CD#98-135.

Is the gap between claim and reality at the bachelor's level something to worry about? For the NLS-72, a decade earlier, this gap was in the 6 percent range (Adelman, 1994). The increase is not statistically significant, but in both cohorts there are significant differences by race and SES, and under those circumstances, the transcripts must be the default.

Since the primary variables in Parts 1 and 2 of this monograph are pre-collegiate, it might also be helpful to ponder the differences between student accounts of grades and course-taking in high school and the evidence of their high school transcripts. Table 2 is extracted from Fetters, Stowe, and Owings's (1984) analysis of this issue in the HS&B/So. It is obvious that we have significant differences in reporting of both grades and course-taking by race, and in course-taking by SES. The case of mathematics course-taking should be particularly troubling to anyone who analyzes pre-college preparation on the basis of student self-reports (in the national data, that includes the annual survey of freshman by the Cooperative Institutional Research Project and the Student Descriptive Questionnaire that accompanies administration of the SAT). Rosenbaum (1980) mapped even greater variances than these in the NLS-72. In two successive cohorts, then, students have been consistent in claiming more coursework than their records show.

Table 2.—Discrepancies between student reports of grades and amount of coursework in high school, by selected student demographic characteristics

							SES Composite		
	All	Men	Wom	White	Black	Latino	Low	Med	High
GPA									
Student	2.84	2.71	2.96	2.91	2.62	2.57	2.64	2.85	3.07
Transcript	2.62	2.51	2.73	2.71	2.31	2.39	2.44	2.63	2.84
Bias	.22	.20	.23	.20	.31	.18	.20	.22	.23
Semesters of Mathematics, Grades 10-12									
Student	4.15	4.31	4.02	4.15	4.50	3.97	3.68	4.07	4.76
Transcript	3.07	3.17	3.03	3.27	2.65	2.39	2.27	3.03	4.02
Bias	1.08	1.14	.99	.88	1.85	1.58	1.41	1.04	.74
Semesters of Science, Grades 10-12									
Student	3.43	3.58	3.30	3.47	3.46	3.13	2.92	3.26	4.09
Transcript	2.87	2.99	2.78	3.00	2.59	2.33	2.29	2.78	3.66
Bias	.56	.59	.52	.47	.87	.80	.63	.48	.43

Note: Adapted from Fetters, Stowe, and Owings (1984), Tables A.4 and A.5, pp. 44-45.

And yet student self-reports continue to be the principal sources of information invoked in the mass of studies on the determinants of college retention and completion. To date, the research community has proven itself intimidated by the richness and power of the details that lie in transcript records. For example, much of the literature on college access was driven by a concern with tracking in secondary schools, and hence collapses the entire range of a student's high school academic background into the dichotomous variable, "academic/non-academic curriculum"—or, sometimes, the trichotomized academic/general/ vocational heuristic for curriculum—thus ignoring some of the most important variations that occur under those umbrellas[10]. It is no wonder that serious consideration of what people study in high school is completely absent from investigations that squeeze the rocks of pre-collegiate "determinants" of college access and persistence, and policy follows suit. People are then surprised when students on putatively "academic" (also known as "college preparatory") tracks wind up in remedial courses in college and/or do not complete degrees.

Table 3.—Selected content and intensity measures for students in high school academic/ college preparatory programs, High School & Beyond/Sophomore cohort

Proportion of Students with . . .	
Highest level of mathematics *less than Algebra 2*	37.2%
Maximum of one year of core laboratory science	38.5
Maximum of one year of foreign language	20.2
8 or fewer credits in core academic subjects	33.2

Notes: (1) Universe consists of all HS&B/So students who graduated from high school and whose program was indicated as "academic." Weighted N=1.44M. (2) The "HSTS" version of the HS&B/So transcripts was used in these analyses.
SOURCE: National Center for Education Statistics: High School & Beyond/ Sophomore cohort; NCES CD#98-135.

In the mid-1980s Alexander and his colleagues began to study curriculum effects with this more empirical flavor, realizing that there was a compelling reason to move away from the dichotomous presentation of high school curriculum, particularly in light of the background research for the National Commission on Excellence in Education (e.g., Adelman, 1983) and its subsequent recommendations for the "new basics" curriculum of *A Nation at Risk*. Using high school transcripts from ETS's Study of Academic Prediction and Growth that tracked students in 15 communities during the period 1961-1969, Alexander and Pallas (1984) found that even among "academic track" high school graduates, only 53 percent met the "new basics" criterion for science, 71 percent did so in mathematics, and a paltry 31 percent matched the mark in foreign languages.

The HS&B/So data allow a more contemporary—and detailed— confirmation. There is an obvious range of intensity and quality in the high school "academic" or "college preparatory" curriculum. Table 3 provides a very simple demonstration. It takes all HS&B/So students for whom an academic curriculum was indicated *by the student's school*, turns to the transcripts, and displays some basic disappointments on the content of that curriculum. These data clearly indicate that some disaggregation of "academic curriculum" is called for. Once again, the more precise and useful data (to guide students onto trajectories leading not merely to college but to degree completion) lie in the details.

Part I:
Cultivating ACRES, the Academic Resources Index

So let us turn to the details. This section of our exploration is essential reading, for in it three indicators of students' high school curriculum and academic performance will be created, then merged to create a master-variable, "academic resources." The acronym for this variable, ACRES, is intended to invoke an agricultural metaphor. It is during the pre-college years that one's academic history is planted and subject to early cultivation.

The Test

The most compliant of the three indicators was that of a senior year test given to nearly all (92.7 percent) of the HS&B/So students. The test can be described as a "mini, enhanced SAT." With a testing-time planned at 68 minutes, the test has core reading, vocabulary, writing and mathematics sections with items drawn from old SATs, plus science and civics sub-tests (Rock, Hilton, Pollack, Ekstrom and Goertz, 1985). The composite score, however, does not include the science and civics sections. Exactly the same test was administered to the HS&B/So group as 10th graders in 1980. While the results are strongly correlated with SAT or ACT scores for the 57 percent of the sample that took either of those tests, they are not psychometrically equatable. All scores for the senior test were set on a percentile scale. Where the senior test score was missing but the student's record included an SAT or ACT score, one could impute a percentile for the senior test, and this procedure was followed for 376 of the 13,477 students whose records exhibit any *de facto* national test scores[11]. Whatever occasionally lumpy effects might result are smoothed when the scale is divided in quintiles. The resulting variable is called TESTQ, or test quintile.

Class Rank and GPA

The construction of the second indicator began with students' high school class rank, by percentile, then quintile. Not all high schools rank students, and one could determine a class rank percent "score" for only 9,082 students out of 13,020 for whom high school transcripts were available. An alternative approach was necessary for the residual group. An academic grade point average[12] was constructed for everyone, and tested against class rank in those cases where both were available. The Pearson correlation was .841. While that is a high number, one is wary of substituting a specific percentile of academic grade point average for a missing class rank percentile because of variances in local grading practices. A larger unit of measurement was necessary to reduce the statistical noise, and quintiles were selected for the task. One could thus substitute an academic GPA quintile for a missing class rank quintile with more than a modicum of confidence. Quintile positions could be determined for nearly all students in the residual group. RANKQ is the result of combining both indicators.[13] It was this variable more than any other that determined the used of quintile presentations for otherwise continuous variables in the analysis.

Academic Curriculum: Intensity

The most complex and important of the variables describes the academic intensity and quality of one's high school curriculum. This construct challenges the five composite variables included in pre-1998 releases of the HS&B/So data that were based on the high school curriculum standards set forth as recommendations in *A Nation at Risk* (1983), the so-called "new basics." These five variables[14] do not reflect a hierarchy, and are impossible to validate. Unfortunately, a new literature has sprung from the use of these variables, so that in place of the old style dichotomy of academic/non-academic we now find "*rigorous* academic" v. "academic" (Akerhielm, Berger, Hooker, and Wise, 1998). "Rigorous," for all its high-sounding educational machismo, is very misleading in this context. "Rigor" applies to standards of performance, which vary widely from course to course and school to school, and no data set can claim to measure it. Academic intensity, on the other hand, is accessible and measurable.

So intensity comes first, and is then modified for quality. For this study, variables covering 15 broad high school subject areas were constructed by standardizing or censoring credit ranges that appeared in the high school transcript files and with reference to state standards for high school graduation (Medrich, Brown and Henke, 1992). For science, two variables were formed: one covering core laboratory science only (biology, chemistry, and physics) and one for all science credits. This distinction was particularly necessary because the transcripts from some high schools did not specify the field of science or offered a "unified" science curriculum. For mathematics, four variables were created: all high school mathematics credits, remedial mathematics units, net mathematics units (all minus remedial), and HIGHMATH, a variable indicating the highest level of mathematics reached by the student in high school. HIGHMATH proved to be an extremely powerful construct, and its position in the subsequent forging of the Academic Resources model was suggested both by Pallas and Alexander (1983), who found some of the elements of the variable on the high school transcripts of the Academic Growth Study[15]; and by Kanarek (1989), who found mathematics-*related* variables (the SAT-Q, an algebra test that was part of the New Jersey Basic Skills battery, total number of years of high school math, the student's self-reported most recent grade in high school math, and the student's self-reported rating of mathematics ability) to contribute significantly to a five-year graduation rate.

The second step in this process was to examine the credit distributions in six core curriculum areas: English, mathematics, laboratory science and total science, history, social studies, and foreign languages, and to cluster them so that five distinct levels of intensity could be discerned. This inductive approach is very different from that reflected in the old "new basics" variables. Table 4 highlights the differences between the highest value of our "academic intensity" indicator and that of the most "rigorous" version of the new basics curriculum configuration. By excluding remedial courses in basic skills and by using a core laboratory science benchmark, we begin to include quality criteria in addition to intensity (Alexander and Pallas [1984] also excluded remedial courses in constructing an empirical counterpart to the "new basics"). The "new basics" variables do not, and the only other major attempt in the literature to set up a detailed curriculum index to be used in multivariate analyses of student-reported "years of postsecondary education" (Altonji, 1994) drew on the NLS-72 high school records, which are *not* transcripts and cannot provide any details concerning the types or levels of mathematics or science[16].

Table 4.—High school units at the highest level of the "academic intensity" variable versus those of the "most rigorous" New Basics variable developed for pre-1998 releases of the HS&B/So data base

	Academic Intensity	"Most Rigorous" New Basics
Units of English	3.75+	4.0+
No Remedial English	Yes[17]	No
Units of Mathematics	3.75+	3.0+
No Remedial Math	Yes	No
All Science Units	(2.5)*	3.0+
Core Lab Science Units	2.0+	---
Social Sci/History	2.0+	3.0+
Foreign Languages	2.0+	2.0+
Computer Science	---	0.5
TOTAL:	13.5+ or (14.0+)*	15.5+

*Only if core lab science units totaled less than 2.0

Why does the highest value of academic intensity use 3.75 units of English and mathematics as a threshold—instead of the 4 unit criterion of the new basics version? The "new basics" variables were constructed from an externally-dictated blueprint. The selection of >3.74 Carnegie unit equivalents, on the other hand, was based on empirical clusters of credits on transcript records from different kinds of high schools with different calendar and credit systems, and in accordance with state requirements. The total of 13.5 or 14.0 academic, non-remedial credits for the highest value of academic intensity and 15.5 credits for the version of "new basics" used here should be compared to a national average of 14.2 academic credits (the remedial and core science strictures not included) for the high school class of 1982 (Tuma, Gifford, Horn, and Hoachlander, 1989). While these bottom lines fall in a fairly narrow range, the constructs are different.

The curriculum intensity variable was modified for quality, adding gradations to each of its five levels for the number of advanced placement courses (0, 1, and > 1), highest level of mathematics reached in high school (+1 for trigonometry or higher, 0 for Algebra 2, and -1 for less than Algebra 2), and subtracting for any case where mathematics course work was largely remedial. The enhanced curriculum indicator then has 40 gradations, set out on a scale of equal intervals from 100 to 2.5.[18] At the highest interval, a mark of 100 on the scale, students display, at a minimum, the following contents of their high school portfolios:

- 3.75 or more Carnegie units of mathematics, with no remedial math
- highest level of math at trigonometry or higher
- 3.75 or more Carnegie units of English, with no remedial courses
- 2.0 or more units of core laboratory science or 2.5 or more units of all science
- 2.0 or more units of foreign language
- 2.0 or more Carnegie units of history or 1.0 unit of history and 1.0 unit of either civics or other social studies
- More than 1 Advanced Placement course

At each of the 40 marks on the interval scale of the enhanced curriculum indicator, one will find a similar richness of curricular description (see Appendix C). The reader should know that this is what lies behind the quintile version of student positions on this scale. It is extremely important for the analysis, conclusions, and recommendations of this study to recognize that the enhanced curriculum indicator is a *criterion* variable: unlike test scores or class rank scales, it sets absolute goals of content, not relative measures of performance. Theoretically, everybody can reach the highest interval of curriculum intensity and quality.

All three component variables were set out in quintiles. To obtain an initial rough estimate of their relative strength in relation to bachelor's degree completion, as well as to see the size of Ns and weighted Ns available for correlations and multivariate analyses, table 5 takes four groups of students and presents a simple cross-tab for each. To be included in these calculations, the student's record had to include *all* of the following: high school transcripts with in-scope credit totals, class rank/academic GPA, senior year test score, and 1993 degree status from the postsecondary transcript file.

Table 5.—Percent of HS&B/So students completing bachelor's degrees, by quintile of performance on three component variables of "academic resources."

	QUINTILES				
	High	2nd	3rd	4th	Low
1) Base Group N=10470; Weighted N=2.83M					
H.S. Curriculum	70	44	19	5*	3*
12th Grade Test	64	37	16	8	3
Rank/GPA	59	36	19	7*	4*
2) On-Time HS Grads Only N=9635; Weighted N=2.61M					
H.S. Curriculum	70	45	19	5*	3*
12th Grade Test	64	38	17	8	3
Rank/GPA	59	37	20	8*	4*
3) On-Time HS Grads with SES Data; N=8819; Weighted N=2.31M					
H.S. Curriculum	69	45	19	6*	3*
12th Grade Test	64	38	17	9	3
Rank/GPA	59	37	21	9	5
4) On-Time HS Grads with SES Data Who Attended College at Any Time; N=6868; Weighted N=1.82M					
H.S. Curriculum	72	49	25	10*	6*
12th Grade Test	67	43	23	14	7
Rank/GPA	64	42	28	14	9

NOTES: (1) The difference between any pair of estimates in a row is statistically significant at $p \leq .05$ except for those indicated by asterisks. (2) Senior Year Weight used for groups 1-3; Postsecondary Transcript Weight #1 used for group 4. **SOURCE**: National Center for Education Statistics: High School & Beyond/Sophomore cohort restricted file, NCES CD#98-135.

What we see in table 5 is that, no matter how one cuts the population, roughly the same proportions earn bachelor's degrees by quintile of the three component resource measures. Furthermore, academic intensity of high school curriculum emerges as the strongest of the three, followed by 12th grade test score, and class rank/academic GPA. It is not surprising, too, that degree completion rates are higher, across all quintiles, when the population is restricted to those who actually entered postsecondary education.

HIGHMATH: Getting Beyond Algebra 2

Of all the components of curriculum intensity and quality, none has such an obvious and powerful relationship to ultimate completion of degrees as the highest level of mathematics one studies in high school. This is a very critical equity issue because not all high schools can offer their students the opportunity to learn the higher levels of mathematics that propel people toward degrees—no matter what their eventual major field of study. If we are serious about preparing students not merely to enter higher education (access) but to complete degrees, then the lesson of what I call the "math ladder" should be heeded, and we will talk more about this issue in the conclusion of this monograph.

Table 6 uses a logistic regression model to illustrate the strong impact of opportunity to learn. The five-rung ladder consists of calculus, pre-calculus, trigonometry, Algebra 2, and less-than-Algebra 2. This HIGHMATH variable was controlled by socioeconomic status, and its analytic unit is an odds ratio, i.e. "the number by which we would multiply the odds [of completing a bachelor's degree] . . . for each one unit increase in the independent variable" (Menard, 1995, p. 49). In this case, the ladder says that, in the High School & Beyond/ Sophomore cohort, for each rung of HIGHMATH climbed, the odds of completing a bachelor's degree increased by a factor of 2.59 to 1. Each rung up the SES quintile ladder (to match the 5 step math ladder), in contrast, increased the odds by a mere 1.68 to 1. To be sure, there are only two variables in this model, but even so, math solidly trounced SES! HIGHMATH correlates with SES at .3095; with bachelor's degree completion at .5103 (see table 7), so the story told by the logistic model is supported by a related measure.

And the precise point at which opportunity to learn makes the greatest difference in long-term degree completion occurs at the first step *beyond* Algebra 2, whether trigonometry or pre-calculus. To be sure, some Algebra 2 courses in high school include trigonometry, but the preponderance of evidence for the period in which the HS&B/So students went to high school suggests that most trigonometry classes were discrete and distinctly labeled. Notice that in the 7-step account of the math ladder, the odds ratio "versus everybody" dips below 1 at the line between Algebra 2 and geometry, with the Beta value going negative at that point. For a moment, it appears that Algebra 2 is the significant cut-point. But when we move to consider sequential odds ratios ("odds ratio versus those below the referent rung"), the line between Algebra 2 and Geometry becomes muddled and the difference not statistically significant, while that between Algebra 2 and trigonometry is a clear break.

If we asked simply what percentage of students at each rung on the math ladder earned a bachelor's degree, the largest leap also takes place between Algebra 2 and trigonometry: a nearly 23 percent increase among all high school graduates, and a 21 percent increase among those whose who continued on to postsecondary education. The empirical account of degree completion, then, reinforces the "speculative" account of odds ratio relationships.

Table 6.—The math ladder: odds ratios for earning a bachelor's degree at each rung, controlling for socioeconomic status (SES), High School & Beyond/Sophomore cohort, 1982-1993

Model with 5 rungs:	Odds Ratio for HighMath:	2.59	t=14.1	p≤.001
	Odds Ratio for SES:	1.68	t= 9.2	p≤.001

Model with 7 rungs:

Highest Math Studied in H.S.	odds versus those below the referent rung	odds ratio versus everybody	Beta	Percent of H.S. Grads Earning BA	Percent of College Students Earning BA	% of All
Calculus	9.52	9.52	2.25	79.8	81.6	6.4
Pre-Calc	7.20	6.15	1.82	74.3	75.7	5.9
Trig	5.42	3.83	1.34	62.2	65.1	11.3
Algebra 2	4.15	1.54	0.43	39.5	44.4	28.3
Geometry	4.27	0.69	-0.38	23.1	28.5	17.0
Algebra 1	2.52	0.17	-1.77	7.8	11.9	20.0
Pre-Algebra	N.A.	0.07	-2.61	2.3	5.1	11.1

NOTES: (1) Universe consists of all students for whom highest mathematics in high school could be determined, for whom SES data were available, and for whom 1993 degree status was known. Weighted N for all high school grads=2.6M; weighted N for postsecondary students=1.95M. (2) Column labeled "% of All" adds to 100.0%. (3) Design effect=1.59. SOURCE: National Center for Education Statistics: High School & Beyond/Sophomore cohort.

A decade ago, a similar attempt was made with the same HS&B/So cohort, without the college transcripts but with student reported college access and February, 1986 degree completion/ Senior status as the dependent variable (Pelavin and Kane, 1990). With an incomplete history (3.5 years after high school graduation) and a dependent variable that is far from the desired end of the story for most students, this analysis announced that completing one or more units of geometry was the critical mathematics filter for those who would be college bound. This conclusion was unfortunate. It is not the *number* of credits in a course that counts, rather, as Madigan (1997) demonstrated with high school science course-taking in relation to tested science proficiency, it is the *level* of courses that should be the unit of analysis. In mathematics, Madigan's advice would be gratuitous. It is obvious, in light of a full 11-year history, that high school students who stop their study of mathematics after completing geometry are not very likely to finish college. In support of their contention that geometry is the cornerstone of advancement, Pelavin and Kane (1990) noted that 29 percent of the HS&B/So students who took one or more year of geometry earned a bachelor's degree or had attained Senior standing by the spring of 1986, As table 6 indicates, that 29 percent is as high as it got, even seven years later, and the percentage is well below the bachelor's degree completion rate for those who reached higher *levels* of mathematics in high school.

Mathematics is the only secondary school subject presenting a distinct hierarchy of courses and that is required for graduation in all states. One could not replicate this type of analysis with any other subject. But in the formulation of "academic resources," the mathematics ladder becomes *part* of a larger construct. It helps us refine gradations of intellectual capital accumulation, and adds a quality dimension to curricular intensity. Its value is thus subsumed, and the variable does not stand alone in the multivariate analyses of this study.

Only 23.6 percent of the college-goers in the HS&B/So cohort (and only 18.4 percent of all high school graduates) reached trigonometry or a higher level of mathematics in high school. If moving beyond Algebra 2 is truly a gateway to higher levels of degree completion, then we have a conditional hypothesis: the higher the percentage of high school graduates who reach those levels of mathematics *and* subsequently attend a 4-year college at any time, the higher the overall system college graduation rate. For the more recent NELS-88 cohort (scheduled high school graduating class of 1992), 37.6 percent reached what the taxonomy for that data set calls "advanced levels" of mathematics (Data Analysis System, NELS-88, NCES CD#98-074). While this NELS-88 indicator is not an exact match with HIGHMATH, it is a credible proxy for the direction of student participation and curricular change during the 1980s and early 1990s. At this time, though, we have no idea whether the empirical or speculative analyses of the HS&B/So cohort will be validated by the NELS-88 since we will not be gathering college transcripts with long-term undergraduate histories for the NELS-88 cohort until 2001.

Correlations, Correlations

Now we can turn to the correlation matrix of table 7, in which the components of "academic resources" are set out. Some 10 variables are used in the matrix, employing the universe of the "base group" of students (see table 5).

Alexander and Pallas (1984) would note that the strong correlations with the 12th grade test score (those of curriculum, highest mathematics, and rank/GPA quintile) are well built into students' momentum by 12th grade, and would hypothesize that similar strengths would be revealed with a 10th grade test that was also administered to the HS&B/So sample. They would also suggest that the one curricular area in which the coefficients increase when appropriate curricular variables are added in a linear regression with test scores as dependent variables is mathematics. And Pallas and Alexander (1983) found that not only the amount, but the level and type of mathematics and other quantitative courses taken in high school was a far more powerful predictor of differences in SAT-Q scores of men and women[19]. This validation of the "differential coursework" hypothesis is one of the principal motivations behind my construction of and emphasis on academic intensity and quality variables—but with degree completion, not test scores, as the outcome.

Some of the correlations in table 7 are weak, suggesting that in multivariate analyses, the variables will not add much to the explanatory power of the equations. For example, on-time high school graduation displays no strong relationships with anything else in the matrix. One might follow McCormick (1999), and use on-time high school graduation as a filter for the

Table 7.—Correlations of major pre-college "Academic Resources" variables and high school graduation status, college entry, and bachelor's degree attainment by age 30, High School & Beyond/Sophomore Cohort, 1982-1993

	Curric. Intensity Quintile (ACINQ)	Curric. Quality Quintile (CURRQ)	Senior Test Quintile (TESTQ)	Cl Rank/ Acad GPA Quintile (RANKQ)	Highest Math (5 Levels) (HMATH)	Math Was All Remedial (RMATH)	AP Courses (3 Levels) (APCRS)	On-Time Grad (ONTIM)	Entered Postsec Any Time (PSENT)
ACINQ	1.000	0.924	0.529	0.465	0.658	-0.415	0.358	0.175	0.336
CURRQ	---	1.000	0.595	0.518	0.756	-0.483	0.388	0.189	0.410
TESTQ	---	----	1.000	0.508	0.540	-0.397	0.337	0.115	0.401
RANKQ	---	---	---	1.000	0.471	-0.315	0.310	0.179	0.331
HMATH	---	---	---	---	1.000	-0.305	0.408	0.114	0.328
RMATH	---	---	---	---	---	1.000	-0.166	-0.129	-0.307
APCRS	---	---	---	---	---	---	1.000	0.073*	0.203
ONTIM	---	---	---	---	---	---	---	1.000	0.118
Earned BA by Age 30	0.509	0.541	0.484	0.441	0.510	-0.255	0.316	0.108*	0.395

Notes: (1) All estimates are significant at $p \leq .001$ except those indicated by an asterisk, which are significant at $p \leq .01$. (2) The universe consists of all students whose records contain positive values for all variables in the matrix. Weighted N=2.689M. (3) Design effect=1.64. SOURCE: National Center for Education Statistics: High School & Beyond/Sophomores, NCES CD #98-135.

statistical noise that might result from including late graduates and GED recipients, but our story will suggest that direct entry to higher education, no matter when one graduates, is more important. Others variables exhibit superficially paradoxical relationships. For example, Advanced Placement course-taking (which, like Highest Mathematics, is subsumed in the curriculum quality variable) is more strongly related to degree completion than to mere entry into postsecondary education, even though 85 percent of those who took AP courses continued their education after high school.

The matrix once again suggests that high school curriculum measures hold a stronger relationship to eventual bachelor's degree completion than the other major secondary school performance measures. Table 8 reiterates what we see of this matter in table 7, but with three outcome measures: entering postsecondary education, attainment of *either* an associate's or bachelor's degree, and attainment of a bachelor's—all by age 30. The table drives home the point that performance has less to do with entering college than it does with completing a degree program, but that no matter what the outcome, curriculum intensity and quality holds the strongest relationships with that outcome while class rank/GPA holds comparatively weak relationships. When we move the threshold of attainment from associate's degree to bachelor's, too, only the changes in curriculum correlations are both positive and significant.

While access (entering postsecondary education) is not the topic of this monograph, one notes in table 8 the comparatively high correlation of test scores with that event. The literature is fairly consistent on this finding. Using the NELS-88 longitudinal study, Akerhielm, Berger, Hooker, and Wise (1998), for example, found that low-income/high test score students entered higher education at a 75 percent rate, compared with 64 percent for high-income/low test score students, and this relationship held up in logit regression models. In an age of aggressive recruiting of minority students, who tend to be in the lower income bands (Cabrera and Bernal, 1998), this is a reasonable finding.

Table 8.—Pearson correlations of the major components of "academic resources" with three outcome measures, by age 30: High School & Beyond/Sophomores

	Entering Postsecondary	Associate's or Bachelor's	Bachelor's
Curriculum Intensity & Quality	.410	.520	.541
Highest Mathematics	.328	.478	.510
Curriculum Intensity Only	.336	.492	.509
12th Grade Test Composite	.401	.483	.484
Class Rank/Academic GPA	.331	.447	.441

<u>NOTES</u>: (1) All estimates are significant at $p \leq .001$. (2) The universe consists of all students whose records contain positive values for all variables in the matrix. Weighted N=2.689M. (3) Design effect=1.64. <u>SOURCE</u>: National Center for Education Statistics: High School & Beyond/Sophomore cohort, NCES CD# 98-135.

Construction of the Composite Variable, "Academic Resources"

It is possible to carry each of these indicators forward, separately, into multivariate analyses. But the more complex the multivariate model, the higher the measurement error under those circumstances. The potential of a composite variable thus arises, and these are used often in NCES data sets. Any composite variable involves a statistical trade-off: one obtains a lower measurement error at the cost of greater covariance, that is, the intrinsic relationships among the components. I preferred the lower measurement error.

Once the three component indicators were developed, they were tested, individually and together, in both linear and logistic regression equations using bachelor's degree attainment as the dependent variable. The weighting of the three components was based on the comparative odds ratios in the logistic equation. Each component was then weighted by its comparative contribution. Table 9 displays the basic model for determining those weights for all students whose records included all three of the component variables, who graduated from high school before 1988 and for whom highest degree earned by 1993 is known for sure. Even though 94 percent of the high school graduates in the HS&B/So sample had received their diplomas or equivalencies by the end of 1983, the pre-1988 boundary for high school graduation was chosen for this calculation because given a mean elapsed time to bachelor's degree of 4.74 calendar years in the HS&B/So, students receiving diplomas by the end of 1987 had the chance to enter college and complete a degree by the time the transcripts were gathered in 1993. For anyone graduating from high school after that point, the chances of meeting the 1993 censoring date for earning a bachelor's degree were nil.

Table 9.–Weighting of the three components of Academic Resources based on their comparative odds ratios in a logistic regression with bachelor's degree attainment by age 30 as the dependent variable, High School & Beyond/ Sophomore cohort, 1982-1993

	Estimate (Beta)	s.e.	*t*	p	Odds Ratio	Weight
Intercept	-3.1794	0.1336	14.5			
Curriculum	0.7252	0.0469	9.4	.001	2.15	40.9%
Test Score	0.4687	0.0464	6.2	.001	1.60	30.4
Class Rank/GPA	0.4150	0.0435	5.8	.001	1.51	28.7

Notes: (1) Universe consists of students with positive values for all three component measures plus known degree status as of 1993, and who graduated from high school by 12/31/87. Weighted N=2.689M (2). Standard errors adjusted for design effects. Design effect=1.64.

Table 10 presents a *linear* regression with the same components, not because it is the source of the weighting, but to indicate that an alternative statistical method yields the same general relationships. For the time being and the purposes at hand, the model is very simple. The adjusted R^2 is solid: it says that, in the absence of any other controls, these three components of the academic resources students bring to higher education account for about 35 percent of the variance in bachelor's degree completion (and where degree completion is unknown, it is assumed to be none). The standard errors are tight, and the indicators of significance are robust (the minimum acceptable *t* would be about 2, and all of the *t*s in this equation are much higher).

Table 10.—Basic linear regression model for the components of academic resources, with bachelor's degree attainment as dependent variable, High School & Beyond/ Sophomore Cohort, 1982-1993

	Estimate (Beta)	s.e.	t	p<	Contribution to R^2
Intercept	-0.4108	.0141	18.8		
Curriculum	0.1076	.0050	14.2	.001	.2947
Senior Test	0.0643	.0051	8.3	.01	.0399
Rank/GPA	0.0533	.0049	7.2	.01	.0176
					Adjusted R^2=0.3521

Notes: (1) Universe consists of students with all three measures plus known 1993 highest degree status and restricted to those who graduated from high school or received a high school equivalency by 12/31/87. Unweighted N=10,233. Weighted N=2.689M. (2) Standard errors (s.e.) are adjusted for design effects. (3) Design effect=1.64. Source: National Center for Education Statistics: High School & Beyond/Sophomore cohort, NCES CD#98-135.

At this very raw preliminary stage, and at the suggestion of one of the reviewers of this study, I tried to bring sex, race (a dichotomous variable, with African-Americans/Latinos/American Indians=1), and a socioeconomic status quintile measure into the basic linear regression model. Neither race nor sex met a very generous selection criterion for statistical significance of .20 (the default selection criterion in the software package is .05). SES, on the other hand, not only met the selection criterion but edged out class rank/GPA for third place (out of four) in contribution to the explanatory power of the model. Consider this exercise a precursor, for SES reflects not merely income, but the kind of parental knowledge of what is involved in higher education, in admissions and college choice, in children's occupational and educational plans, and financial planning–all of which contribute to a supportive environment for student persistence (Flint, 1992). These socioeconomic factors, by-products of parental occupation and/or level of education, have been persistently shown to be far more important than race or sex in relation to a child's degree attainment (Burbridge, 1991; Hearn, 1991).

Once a student has a quintile score for a given component, for example, Senior Test, it is multiplied by its relative strength weight. In this method, the scale for Academic Resources is compressed to 1-5, and the sum of the adjusted component weights is the basic "Academic Resources Index." A student in the highest curriculum quintile, the second highest test quintile and the second highest class rank/GPA quintile would have a composite academic resource index of 4.409 (5x.409 + 4x.304 + 4x.287). Index scores such as these were again set out in quintile bands. The final product is the variable called ACRES. It is this variable that we carry forward into multivariate analyses.

This method of creating an academic resources index differs somewhat from its sources of inspiration. Alexander, Pallas, and Holupka (1987), for example, boiled 18 combinations of curriculum, test scores and grades down to four groups. Their formula took a standard deviation above and below the mean for both test score and GPA, and crossed the results with a dichotomous curriculum variable, academic/non-academic. Their extreme groups are thus roughly equivalent in size to the tails of a standard distribution, that is, the top 16 percent and the bottom 16 percent. While this approach carries a strong academic logic, it is not as persuasive in public policy applications as would be an array of standard and fairly transparent intervals such as quintiles, nor does it highlight the kind of curricular details one would assume from a "differential coursework hypothesis." In a similar vein, it would be possible to use a standard distribution of "academic resources index" scores, isolating the tails, and blocking the middle ranges in Standard Deviation Units (see Alexander, Riordan, Fennessey, and Pallas, 1982, and Alexander, Holupka and Pallas, 1987). That strategy, however, again undercuts the very purpose of the curriculum portion of the index, which relies on benchmarks that, theoretically, all students can reach *if they have the opportunity*. The SDU strategy also results in a tripartite division of academic resources[20] that does not match the accessible quintile formulation of the key SES variable.

Does ACRES Work?

Prior to multivariate analysis, there are two ways to illustrate whether the composite variable tells a consistent story against its principal "rival" in this early stage of investigation, socioeconomic status (SES). ACRES and SES, of course, are hardly "rivals": the Pearson correlation between the two variables is .368, a modestly strong relationship. Even so, table 11 provides some clues as to the extent to which Academic Resources can overcome the effects of SES. While there is a linear relationship between both variables and bachelor's degree completion, the curve for Academic Resources in much steeper. The long-term degree completion rate for those in the highest quintile of ACRES is 72.5 percent, 17 percent higher than for those in the highest quintile of SES. Yes, the higher one's initial SES quintile, the stronger one's platform for launching an effort to earn a bachelor's degree, but acquiring academic resources pays off at a higher rate of interest, so to speak. Among those who attended a 4-year college at any time, the ACRES story is more consistent than the SES story.

Table 11.—ACRES versus socioeconomic status in relation to bachelor's degree completion, High School & Beyond/Sophomore cohort, 1982-1993

	Percent Completing Bachelor's Degrees Within Each Quintile				
	High	2nd	3rd	4th	Low
All High School Graduates					
Academic Resources	72.5	44.8	17.3	5.4	2.0
Socioec Status	55.4	32.9	23.2	14.5	7.2
All Who Attended a 4-Year College					
Academic Resources	80.3	64.1*	40.1	25.1	16.7
Socioec Status	72.1	59.6*	55.5*	45.5	35.4

Notes: (1) The first universe of all high school graduates includes students who received a diploma or GED by 12/31/87 for whom SES and ACRES could be determined. Weighted N=2.45M. The second universe restricts the first by attendance at a 4-year college at any time. Weighted N=1.15M. (2) The Senior Year Weight was used. (3) All row and column pair comparisons are significant at $p \leq .05$ except those indicated by asterisks. SOURCE: National Center for Education Statistics: High School & Beyond/Sophomore cohort, NCES CD#98-135.

Another, and more dramatic, way to illustrate these relationships is to present bachelor's degree completion rates within each quintile of Academic Resources, controlling for SES. Table 12 provides a descriptive account for all high school graduates in the HS&B/So who subsequently attended a 4-year college at any time. Why is this a more dramatic account? Because it shows, for example, that students from the *lowest* two SES quintiles who are in the *highest* ACRES quintile earn bachelor's degrees at a higher rate than a majority of the students from the highest SES quintile. It demonstrates that students in the bottom two quintiles of ACRES earn degrees at low rate *no matter what their SES standing*.

This is a temporary judgment in the unfolding of this study. It is temporary because the universe (in table 11, in particular) has been divided in a rather stark manner: all high school graduates v. those who went on to attend a 4-year college at some time. There are students who aspire to 4-year degrees but never attend 4-year colleges, and their background characteristics may be sufficiently different from others to warrant a recalibration of the balances among SES, Academic Resources, and degree attainment. The most productive moment to confront this issue is at the beginning of the multivariate analysis in Part IV.

Table 12.—Bachelor's degree completion rates by quintile of ACRES, controlling for socioeconomic status, High School & Beyond/Sophomore cohort, 1982-1993

	Percent Completing Bachelor's Degree Within Each ACRES Quintile				
	High	2nd	3rd	4th	Low
Socioeconomic Status Quintile					
Highest	85.9	75.4	51.2	28.1*	12.8*
2nd	79.2#	59.5#	42.7#	20.7*	20.1*
3rd	78.6#	58.7#	38.4#	29.5*	17.4*
4th	66.0##	60.6#	24.5*	26.5*	14.6*
Lowest	62.2##	42.1	28.9*	20.4*	20.8*

Notes: (1) Universe consists of all students who received high school diplomas or equivalents by 12/31/87, for whom socioeconomic status and ACRES could be determined, and who attended a 4-year college at any time. Weighted N=1.15M. (2) *=row and column comparisons that are *not* statistically significant. (3) # or ##=column pair estimates that are *not* statistically significant. (4) All other comparisons of estimates are significant at $p \leq .05$. SOURCE: National Center for Education Statistics: High School & Beyond/Sophomore cohort, CD#98-135.

Summary

As a composite variable heavily weighted by intensity and quality of high school curriculum, academic resources (ACRES) is a valid and viable construct to represent the intellectual capital brought by students to the threshold of postsecondary education, and can be profitably carried forward into analyses of postsecondary careers. For the High School & Beyond/Sophomore cohort, the components of ACRES reflect critical details of secondary school performance, not deceiving generalized dichotomies. ACRES is a heuristic in which we can have great confidence, far more than in a label such as "college qualified" that Berkner and Chavez (1997) constructed with reference to the pre-collegiate performance of students who *enter* 4-year colleges[21]. How much it will contribute to our understanding of bachelor's degree completion when other background variables and college experiences are brought into play remains to be seen. It is precisely because our dependent variable is long-term (by age 30) bachelor's degree completion for a conventional age cohort that we now pause to consider what the completion issue entails and what other variables traditionally used in analyses of educational attainment must be either discarded or reconstructed.

-II-
Touchstone Variables:
Persistence, Completion, Aspirations, Parents, and Parenthood

"Academic Resources" moves forward in our story onto a field with other variables that historically have served as touchstones in analyses of educational careers. Some of these touchstones involve the characteristics and environments of the institutions attended, and these are addressed in Part III. In this section, though, we are concerned with sharpening the focus on our dependent variable, degree completion, and with the sources and shape of other major background variables that researchers, policy-makers, and interpreters of trends in education tend to accept as if they were holy writ. A more secular framework of judgment is called for.

The principle points to be made are:

(1) Persistence is not degree completion. In fact, the temporal measurement of persistence is deceiving, and should be replaced by credit-generation thresholds. Completion transcends persistence.

(2) If persistence is a sub-set of completion and completion is a more compelling notion, so is the relationship between some components of socioeconomic status (SES) and the larger construct. We know we have difficulty with student reports of family income, one of the principal components of SES. But student reports of parents' highest level of education turn out to be unreliable as well. The lesson is to stick with the umbrella of SES and not to use its pieces separately.

(3) It is remarkable that while research on persistence and degree completion pays a great deal of attention to parents, it rarely marks parenthood as a characteristic of students themselves. Parenthood, however, is not as easy a variable to define as one might think.

(4) Historically, research on persistence and degree completion has paid considerable attention to educational aspirations, a notion usually elicited on the basis of one or two questions in a survey. The wording of these questions, however, elicits a statement of "expectations," a very different construct from aspirations. The variable is too important for superficial construction. The HS&B/So data base allows us to reconstruct it in a more meaningful way, as "anticipations."

Persistence and Attrition

Most previous research employing pre-collegiate input variables has focused on the determinants or correlates of access or enrollment (for a summary of this literature, see Baker and Velez, 1996).[22] There are two sets of literature that take us beyond this access threshold. The first uses "persistence" as the dependent variable, but seems less concerned with persistence than with attrition (DesJardins, Ahlburg, and McCall, 1997; McConnell-Castle, 1993).

Tinto (1975, 1982, 1993) set the terms in this literature, with attention to social and academic integration factors in students' experience; and the research deriving from Tinto's work is vast. Bean (1980, 1983) and Metzger and Bean (1987) offered a different kind of analysis of attrition based on organizational process theory and the phenomenon of personnel turnover within organizations, and, as a consequence, have been the only researchers to emphasize features of institutional location (distance from the student's home, and in-state attendance) as factors in student careers[23].

Both lines of research treat persistence or attrition as a phenomenon existing *within a given institution* (Cabrera, Nora, and Castaneda, 1993), sometimes illustrating the problem of single-institution studies in which the student population lies considerably above academic performance means (e.g. Bean, 1980; Bean, 1982). If a community college population is the subject, the studies also illustrate the problem of defining persistence and attrition in terms of *intent* to transfer (see, for example, Nora and Rendon, 1990), and the problem of influences of family and significant others on a predominantly commuting population (Nora, Attinasi, and Matonak, 1990)[24]. Both approaches, as Cabrera, Castaneda, Nora and Hengstler (1992) point out, regard persistence and attrition as the result of interactions among precollege characteristics, college environments, and adjustments to college. Both use college grades and continuation beyond the first year of study as outcome variables (see, for example, Pascarella and Terenzini, 1980; St. John, Kirshstein, and Noell, 1991). Both research traditions place an extraordinary emphasis on psychological variables: intentions, attitudes, influences, commitments, perceptions (see, e.g. Cabrera, Castaneda, Nora and Hengstler [1992]). These variables unfortunately refer to realities that lie beyond the control of those who can best steer students toward degree completion.

"Persistence" itself, as commonly used in the research literature, may not be a convincing dependent variable. We have no idea, for example, what a one-year "persistence rate" means. Does the student who arrives at something called "year 2" get there with 11 credits, 19 credits, 32 credits? In the HS&B/So, of all students who entered college directly from high school in 1982 *and* persisted through 1983, 17 percent arrived at "year 2" of their college careers in the summer or fall of 1983 with less than 20 credits, and another 12 percent earned more than 20 credits but with three or more remedial courses. In other words, nearly three out of ten "persisters" evidenced less than what one would call "sophomore standing." Table 13 displays some of the variants in this portrait. It is no surprise that students with low credit accumulation in the first year who also show "temporal persistence" are far less likely to earn degrees by age 30. Without credit accumulation information, structural equations with "persistence" as an outcome are very deceiving, and are apt to overstate the influence of affective factors as opposed to academic achievement.

Unlike "persistence," the completion of a bachelor's degree is a censoring event, the culmination of years of preparation and effort. Momentum toward that event is not necessarily measured in years, however, rather in accrual of the currency of the degree (McCormick, 1999). And the more currency acquired, the higher the odds of completing. To launch the

Table 13.—The fallacy of temporal persistence: percent of 1982 HS&B/So college entrants who "persisted" to the academic year 1983-1984, by credits earned in the first year, number of remedial courses, and degree completion

Credits Earned:	0-12	13-19	20-28	>28	% of All:
ALL	7.6%	9.3%	31.2%	51.9%	
By Number of Remedial Courses					
None	LOW N	5.2*	28.6	62.3	48.5
One	8.1*	8.4*	30.9	52.6	21.8
Two	12.0*	13.4*	33.9	40.7	12.0
Three or More	14.1*	19.0	36.6	30.3	17.7
By Highest Degree Earned					
None	19.4	19.9	36.5	24.2	29.7
Associate's	LOW N	10.4*	26.4	56.1	10.6
Bachelor's or Higher	1.9*	3.9*	29.4	64.8	59.8

Notes: (1) Universe consists of all on-time high school graduates who entered postsecondary education between June and December, 1982 and who were also enrolled during the Academic Year beginning July 1, 1983. (2) Weighted N=1.2M; (3)*Low-N cells with no statistical significance. (4) Rows for credits earned may not add to 100.0% due to rounding and Low-N cells. SOURCE: National Center for Education Statistics: High School & Beyond/Sophomore cohort, CD #98-135.

measure of completion rates from credit accumulation thresholds provides a fairly strong guidance. Table 14 expands the framework suggested by table 13, and offers some hints of what might happen if one set post-matriculation credit thresholds in regression equations (ordinary least squares or logistic) in which bachelor's degree completion is the dependent variable. Table 14 also marks students who entered higher education within six months of *on time* graduation from high school in 1982. The reader will note that these "direct" students completed bachelor's degrees at a rate higher than that for all students.

This study takes the position that bachelor's degree attainment rates should be measured only for those people who actually attended a 4-year college at some time. Students can tell us on surveys that they intend to earn a bachelor's degree, but if they never set foot in a 4-year college long enough to generate a record, they made no attempt to do so, and it is misleading to include them in the denominator of potential degree recipients. Long-term bachelor's degree attainment rates jump dramatically as soon as one confines the universe to those who have attended 4-year colleges, and exhibit impressive gains with each ratcheting up of the credit threshold. The effect of no delay of entry to higher education also plays a positive role, but the strength of that role diminishes with credit accumulation. If we wished to push the conditions for degree attainment to near-maximum levels, we would also limit the institution of first attendance to a 4-year college, insist that more than 30 credits be earned from 4-year colleges, and add the direct entry criterion. The long-term (eleven year) degree completion rate under those conditions is 79.1 percent, and is higher, still, if the first institution of attendance was highly selective and if the student attended only one college.

Table 14.—Percent of students completing degrees, by credit-generation thresholds, High School & Beyond/Sophomore cohort, 1982-1993

	All Postsecondary Students			**Attended a 4-Year College At Any Time**					
Credit				**>0**		**>10**		**>30**	
Threshold:	≥0	≥10	≥30	All	DIR	All	DIR	All	DIR
Highest Degree by Age 30:									
None	50.7	45.1	32.7	31.5	26.8	29.5	25.6	22.7*	20.1*
Associate's	9.3	10.4	12.6	5.6*	5.5*	5.8*	5.6*	6.3*	6.0*
Bachelor's	40.0	44.5	54.8	62.9	67.7	64.7	68.8	71.0	73.9
Weighted N	1.95	1.84	1.48	1.31	1.12	1.27	1.11	1.16	1.03

Notes: (1) DIR=Direct Entry. (2) Universe=all students for whom transcripts were received. (3) Weighted N is in millions. (4) Columns may not add to 100.0% due to rounding. (5) Paired comparisons (All v. DIR) are significant at $p \leq .05$ except those asterisked. SOURCE: National Center for Education Statistics: High School & Beyond/Sophomore cohort, CD# 98-135.

In a similar analysis, McCormick (1999) formulated a scaled-down version of the same database used in this study, excluding about a third of the students in the HS&B/so who ever attended a 4-year college as well as certain types of credits[25]. McCormick was more interested in the *pace* at which students accumulate credits, but his conclusions about degree completion rates mirror those indicated above, and, with some minor variations, those displayed in table 14. On balance, McCormick's degree-completion rates are higher because of the exclusions, but the relationships between credit-generation thresholds and completion rates are the same, a reassuring consequence of using the same transcript-based source.

Completion

As noted above, there are very few national studies across the entire literature on persistence and attrition that hold the completion of a degree to be the sole and/or most prominent dependent variable. Researchers often tout the importance of completion studies and then tell us why they cannot conduct such studies with a full sample of students. For example, Hauser and Anderson (1991) beg off by noting that "changes in the timing and intensity of college attendance have made it more difficult to measure and compare rates of college completion in recent years" (p. 275), though, in fact, the more complex attendance patterns have become, the richer the analysis (see Part III below). Chaney and Farris (1991) conducted an analysis of completion, but confined it by an institutional-entry cohort restricted to full-time students, capped the measurement period at 6 years, and reminded us that once students leave the institution they are not tracked, therefore "institution-specific retention rates may significantly understate retention within higher education as a whole" (p. 5). Astin, Tsui and Avalos (1996) offer a similar disclaimer, though they should be credited for seeking evidence of degree completion from institutional registrars as opposed to relying on student self-reports. Their 9-year follow-up to the entering freshman class of 1985 is a commendable advance on the time frames of completion studies grounded in data bases other than NCES longitudinal studies, but resulted in a low response-rate from the registrars[26]. What we inherit from all these talented researchers, then, are partial portraits of completion.

Carroll established a model for national analysis using the history of the High School & Beyond/Senior Cohort (high school class of 1980, followed to 1986) to define a "persistence track" that ends with bachelor's degree attainment. Carroll's principal variables are those of immediate entry to higher education following high school graduation, full-time enrollment, stop-out behavior, and transfer. His strategy was to set an "optimal" model, and then to demonstrate how year-to-year deviations from the model resulted in either no degree or a time-to-degree beyond the classical 4-year norm. Carroll's analysis was the first to take account of attendance patterns in a sophisticated way. It is a descriptive portrait of dynamic flows without pretending to predict or assert deterministic relationships. While Carroll cites pre-college variables, he uses them minimally: his focus is on the "persistence track."

Carroll used survey data (not transcripts) and hews to a discrete time hazard model that has been shown to be a productive approach to the understanding of attrition (DesJardins, Ahlburg, and McCall, 1997). The data source, however, creates a particular problem in student self-reports of part-time status. It is difficult to determine the fraction of Carroll's universe that either started part-time or shifted to part-time status, but, in the HS&B/Sophomore cohort, the proportion of students who told us they were enrolled part-time at any time in their college careers was very low: 15.2 percent[27]. This fraction is so far below what institutions report on the federal Integrated Postsecondary Education Data System (IPEDS) surveys as to cast students' judgments in severe doubt. In 1984, the modal year of postsecondary attendance for the HS&B/So, institutions reported that 40.2 percent of their undergraduate enrollments were part-time (Snyder, 1987, p. 129). A decade later, we witness the same type of discrepancy: in the fall semester of 1995, 41.6 percent of undergraduates were reported as part-time students

(*Digest of Education Statistics, 1997*, table 175, p. 185); in the Beginning Postsecondary Students Longitudinal Study of 1989-1994, only 26.5 percent of students ever reported being enrolled part-time as undergraduates[28]. Sometimes, one is unsure whether students know what "part-time" means.

Carroll's two tracks have strong boundaries. The moment one violates the boundaries, one is declared off-track and no longer persisting. These "violations" include reverse transfer (4-year to 2-year), a shift to part-time status, and stop-out (defined as a gap in attendance of 4 non-summer months). In combination with delayed entry and enrollment in a 2-year school or in a nondegree-granting postsecondary institution, these behaviors are generally acknowledged as a drag on attainment (Pascarella and Terenzini, 1991; Hearn, 1992).

Berkner, Cuccaro-Alamin and McCormick (1996) borrow and extend Carroll's scheme. Their data were drawn from the Beginning Postsecondary Students Study of 1989-1994. These data are not exactly comparable to those from age-cohort studies such as the HS&B/So: nearly 27 percent of the participants were 20 years or older in their first year (1989-90) and nearly 10 percent were at least 30 years old. Dates of high school graduation are spread over two decades, and very little is known about the pre-college backgrounds of participants. The authors configure a set of characteristics of students at the moment of first entry to higher education (the "event" that defines the cohort) that are dominated by household, employment, and dependency variables (number of children, single parent status, full-time employee,[29] and independent status as defined in federal financial aid regulations). These characteristics, which are highly correlated with age at entry, along with the *quality* of entry in terms of delayed enrollment and part-time status, they say, form a collection of "risk factors" that work against persistence and completion of *any* credential, certificates and associate's degrees included.

Berkner, Cuccaro-Alamin, and McCormick's study is valuable for its methods, and, advancing on Carroll's rules for defining a "persistence track," the attention it pays to attendance patterns, particularly in the combination of transfer and continuous/noncontinuous enrollment. The database allows the analyst to construct at least a rough approximation of the kind of academic and social integration indices used by Tinto and Astin throughout their work, something we cannot construct in the HS&B/So. The dependent variables included in Berkner, Cuccaro-Alamin, and McCormick's model of system persistence/attainment include credential award (certificate, associate's degree, or bachelor's degree), and final enrollment status (still enrolled/no longer enrolled). "Internal persisters" are distinguished from "transfer persisters." Based on the institution of first attendance, the basic "persistence history" of this cohort can be summarized as follows:

- the first institution of attendance "retained" 43.3 percent of its students;

- but when one adds the *students* who transferred from the first institution and who either earned a credential (of any kind) or were still enrolled at another institution at the end of the tracking period, the result is a 63.1 percent *system* attainment/persistence rate;

- transfer students also exhibit the highest composite rate of attainment/persistence (68.8 percent) of any division of the BPS90 universe.

These data provide hints that are supported elsewhere in the literature (see, for example, different kinds of evidence from Lee, Mackie-Lewis, and Marks, 1993; Blumberg, Lavin, Lerer, and Kovath, 1997; McCormick, 1997; Adelman, 1998) that the very act of transfer embodies an intensity of commitment to higher education that results in degree completion rates equivalent to—if not higher than—those of students who do not transfer. It is no wonder that, as we will see, multi-institutional attendance (with or without formal transfer) is not a drag on degree completion. Institutions may "retain," but students "persist."

Aspirations, Expectations or Anticipations?

One of the most persistent variables in the persistence literature measures educational aspirations. In the NCES longitudinal studies, we ask students about their vision of their future education at each survey occasion, and invest the responses with a great deal of explanatory power in terms of subsequent behavior (e.g. St. John and Noell, 1989; Pelavin and Kane, 1990; Hauser and Anderson, 1991). But we must be careful about such evidence, pay close attention to the wording of the questions, and select those occasions on which the evidence is most persuasive. If, as in the NLS-72 survey, one asks a senior in high school whether he/she "would like" to attend college and enters the responses in a structural equation with college entrance as the dependent variable, it is like looking at a sky filled with dark clouds and predicting rain (Alexander and Cook, 1979).[30] After two years of college, asking whether someone "plans" to complete a bachelor's degree is a similar exercise. On either occasion, there is enough momentum in students' educational histories so that they know the likely outcomes on the path they are taking. In other words, on these occasions we are no longer dealing with "aspirations," rather with "expectations." The resulting high correlation between expectations and access (Akerhielm, Berger, Hooker, and Wise, 1998) is highly artificial. Family expectations are part of this momentum, usually a reflection of SES. Asking whether a student's expectations agree with those of parents is more likely to produce concurrence among higher SES students than lower SES students. In the 1982 (senior year) HS&B/So survey, 30 percent of students in the lowest SES quintile did not report or did not know their parents' expectations for their postsecondary education--versus 12 percent for students from the highest SES quintile. It is for this reason, in particular, that some analysts place the "aspirations" variable far down the line in regression models (see, e.g. St. John, 1991, who noted that "it is appropriate to consider its influence only after . . . other factors have been considered," p. 145).

Students are asked not only about aspirations (in Hauser and Anderson's [1991] phrasing, "desired outcomes that are not limited by constraints on resources," p. 270), but also about *plans*, a more realistic assessment of future action. Students who say they aspire to or expect to earn a bachelor's degree, but also indicate that they plan to spend most of the year following high school graduation in an apprenticeship program illustrate one form of the difference between aspirations and plans: the aspiration is a generalized guide, but the plan to execute it is elusive. The extent of such out-of-scope choices in the HS&B/So is notable: 11.2 percent of those who expected to earn graduate degrees and 15.2 percent of those who expected to earn bachelor's degrees planned to engage in activities during the year following high school graduation that would lead them nowhere near the paths to those goals.

How, then do we best measure the consistency and strength of pre-college educational expectations in the HS&B/So? The surveys asked six questions in both the 10th and 12th grades for this cohort that help up arrive at a more sophisticated notion. When we look closely at the wording of these questions, we realize that they do *not* address aspirations at all. They reference neither constraints not goals. Some express expectations. Others ask for concrete plans. Still others ask for projected affective states. Together, they map a complex set of what we might call "anticipations." These questions, spread out over the survey forms, are:

- As things stand now, how far in school do you think you will get?
- What is the minimum level of education with which you would be satisfied?
- What activity most likely will take the largest share of your time in the year after you leave high school?
- Do you plan to go to college at some time in the future (directly from high school, after a delay, don't know, no)?
- If you went to college, would it most likely be a 4-year or a 2-year institution?
- Would you be disappointed if you did not graduate from college?

We have 12 pairs of responses, then, with which to build an "anticipations" variable based on the concepts of consistency and level. The five resulting gradations of the variable are:

1) Consistent expectations for a bachelor's or higher degree. No matter what question is asked, the student had the same answer in grade 10 and grade 12, and no answer departed from a story line of entering a 4-year college directly from high school, earning a bachelor's degree, not accepting less and being disappointed if he/she were not a college graduate. The only change allowed is a decrease in expectations from a graduate degree to bachelor's.

2) Either raised expectations between grades 10 and 12 to a bachelor's degree or evidenced consistency in some pairs of questions about bachelor's degree-oriented behaviors (1980/1982) but not in others.

3) Either lowered expectations from bachelor's to a sub-baccalaureate credential, or indicated bachelor's degree expectations at some time, but undercut indicated expectations with *consistent* "sub-bachelor's" responses to questions about primary activity in the year after high school or plans to attend college in the future. Consistent associate's degree aspirants were included at this level.

4) Lowered expectations from a bachelor's degree to no degree and/or evidenced considerable confusion about future plans. Inconsistent associate's degree aspirants were included at this level.

5) Expected no degree of any kind as both sophomores and seniors in high school or lowered expectations from a sub-baccalaureate credential to no degree. The category also includes a group who said they didn't know the level of education to which they aspired and had no college plans for the year following high school.

How strong is this variable? Table 15 indicates what happened to the populations at each level of "anticipation" in terms of degree attainment. Among all postsecondary students there is a very clear and dramatic linear relationship between bachelor's degree attainment and the levels of anticipation. As soon as one drops below the "bachelor's-consistent" level, the bottom falls out on long-term degree completion. But among those who attended a 4-year college at any time, the linear relationship is not as dramatic, and only the small group of those who never aspired to the bachelor's degree (despite their attendance at a 4-year school) evidences very low degree completion rates.

Table 15.—Degree anticipations and highest degree earned by age 30 in the HS&B/So: percent of students completing degrees at each level of anticipation.

	All Postsecondary Students				All Who Ever Attended a 4-Year College			
ASPIRE	No Degr	Assoc	Bach	% of All	No Degr	Assoc	Bach	% of All
Bachelor's Consistent	31.5	4.0	65.5	**33.9**	25.1	3.3	71.6	**53.4**
Increased to Bachelor's	55.3	11.7	33.0	**28.6**	36.2	7.4	56.4	**28.9**
Associate's Consistent or Reduced from Bachelor's	67.2	16.7	16.1	**17.4**	43.6	12.8*	43.6	**11.1**
Certificate or Associate's: Inconsistent	88.6*	6.3*	5.1*	**8.5**	72.2*	5.2	22.6*	**3.4**
No Degree or Never Knew	89.0*	7.6*	3.4*	**11.5**	69.0*	10.1*	20.9*	**3.3**

NOTES: (1) The universes consist of all students for whom "anticipations" could be determined from survey responses in both the 10th and 12th grades. For all postsecondary students, the Weighted N=2.38M; for those who attended a 4-year college at any time, the Weighted N=1.37M. (2) Column pairs are significant at $p \leq .05$ *except* those marked by asterisks. SOURCE: National Center for Education Statistics: High School & Beyond/Sophomore Cohort, NCES CD#98-135.

The "anticipations" variable has five values but is not like the quintile-formatted variables used in this analysis: it is not based on a continuous scale such as socioeconomic status (SES) or senior year test percentiles, nor it is based on intervals that yield roughly equivalent quintiles, such as the basic "academic resources" variable (ACRES). The variable calls for a dichotomous reconstruction. Among students who attended a 4-year college at any time, the positive side of a dichotomous variable would be confined to those who exhibited "bachelor's consistent" expectations. The true value of anticipation in student histories lies in its correlation with attending a 4-year college (whether initially or through transfer). But in Part IV of this study, confined to 4-year college students, its position is more complex.

It is worth noting that Morgan (1996), examined changes of the HS&B/So students' responses to the "aspirations questions" (but not the minimum satisfactory, dominant activity, or future plans questions) between grades 10 and 12 and treated them as a dependent *continuous* variable. Morgan found a larger proportion of students (45 percent) who did not change their goals between grades 10 and 12 than is reflected in our account of "anticipations." This is an intriguing approach, but the story it tells ends at the variable[31], not with actual attainment.

Level of Parental Education: a Dubious Investment

Much of the previous research on postsecondary access and attainment employs stepwise regressions in which key components of socioeconomic status are treated as separate entities. Most notable among these components are parents' occupations and level of education, and family income.[32] In nearly all national data sets, these data are reported by students.[33] One can smell the problems with this reporting from a great distance (Fetters, Stowe, and Owings, 1984). Other perceptive analysts (e.g. Mare, 1980) have cited the low reliability of such variables in light of children's changing understanding of their parents' educational history and occupational status, and still others (e.g. Morgan, 1996) have gone to the creative trouble of substituting parents' responses (when available) and restandardizing the SES scale for analyses of NCES longitudinal studies. Indeed, a comparison of student and parent accounts in the NELS-88 dataset (see table 16) canonizes these observations and adjustments.

Late adolescents may know what their parents do for a living, but their idea of the highest level of education attained by their parents leaves something to be desired (Mare, 1980). As high school students, some 29 percent of the HS&B/So participants found various ways to indicate that they did not know their father's highest level of education; though for mothers, the rate was a mere 19 percent. Furthermore, in 8 percent of the cases for those who claimed to know their parents' highest level of education, there was a raw dissonance between reported parental occupation and reported parental education. For example, according to the students, we have lawyers whose highest degree was "some college," and school teachers whose highest level of education was "high school graduate." One can edit some of these cases in the database, but only where the parental occupation indicated requires at least a college education.[34] Other researchers have found that parents' educational levels have little direct effect on success, particularly when compared to pre-college acquisition of knowledge and skills (Mow and Nettles, 1990; Grandy, 1998), so there is some indirect support for the position taken here.

Since the High School & Beyond files include interviews with a sub-sample of parents, we can compare perception to reality. Fetters, Stowe and Owings (1984) confirm the child's tendency to underestimate their parents' levels of education, even when they can classify their parents' occupations accurately. For example, when they were in grade 10, 35.3 percent of the HS&B/So cohort reported their parents had continued their education after high school, whereas 44.4 percent of the parents reported having done so. Among high school seniors, 37.4 percent reported postsecondary education for their parents, compared with 43.2 percent by the parents' account (Fetters, Stowe, and Owings, table A.2, p. 41). These are uncomfortable discrepancies made more uncomfortable by public policies that encourage "first generation college students" to continue their education, let alone by inter-generational social mobility analyses that rely on imperfect second-party accounts of educational attainment (Hearn, 1984; Karen, 1991; Lang, 1992).

Table 16.—Degree of agreement on parents' highest level of education in the NELS-88 longitudinal study.

	Highest Level of Education Attained by Either Parent					
	Didn't Finish H.S.	H.S. Grad	Some College	College Grad	Master's	Ph.D. or Profess.
Parents' Account:	11.4%	21.2%	41.0%	14.0%	8.5%	3.9%
Highest Percentage of Students Agreeing With At Least One Parent:	46.7*	41.6*	72.3	50.6	55.0	63.5

Notes: (1) Universe of students consist of all who answered questions about their parents' highest levels of education in 1992 (and, if not then, in earlier surveys), including those who indicated they did not know but excluding missing cases. Weighted N=3.03M. (2) F3PAQWT, a weight for the NELS88 "parents' file," was used. (3) *Highest degree of agreement was with mother; otherwise, highest degree was with father. (4) Row for parents' account adds to 100.0% SOURCE: National Center for Education Statistics: Data Analysis System, NELS-88.

If the HS&B/So data are not enough to scare one away from using student reports of their parents' highest level of education, perhaps the NELS-88 longitudinal study, which contains a substantial parents' file, will finish the job. Leaving aside the 15.7 percent of the NELS-88 students who would not venture a guess about their parents' schooling, table 16 presents some major indications of conflict in student and parent accounts. In these data, students emerge as more likely to understand that their parents have *attended* college than earned degrees, and seem to have a more definite sense of college attendance and credentials for fathers than for mothers. Of course, it is possible in all these data that the *parents* are inflating their educational attainment, but the conflicts between reported occupation and educational level that we observed in the HS&B/So suggest otherwise.

Parenthood Itself

When we follow a traditional age cohort such as that of the HS&B/So, what Berkner, Cuccaro-Alamin, and McCormick's 1996 study (along with an earlier exploration of attrition among "non-traditional" students in Bean and Metzger, 1985) reminds us to do is to account for change in family status as students move into their early and mid-20s. The variable of choice is parenthood, and the reader will note the importance of this variable in Part IV below. In a correlation matrix with college access as the dependent variable, having a child prior to age 20 ranked third (behind SES and family income)—and well ahead of race, among demographic factors associated with entering postsecondary education, in this case, a negative association. Table 17 excerpts the critical data from the correlation matrix. The NELS-88 data provide continuing confirmation of these relationships. For example, of students in the lowest SES quintile in the NELS-88, 31 percent had children by age 20; and of those who had not entered postsecondary education by age 20, a third were already parents.

Table 17.—Pearson correlations for demographic factors in access to postsecondary education and early child-bearing for the High School & Beyond/Sophomore cohort

	Entered PSE By Age 30	Had Child by Age 20
SES (quintiles)	.3261	-.1501
Family Income	.2255	-.1326
Child by Age 20	-.1998	----
Race (minority=1)	-.1007	.1140
Sex (female=1)	.0601*	.1095

Notes: (1) Universe consists of all HS&B/So students who participated in both the Base Year (1980) and second (1984) follow ups, and whose files contain positive values for all variables in the matrix; (2) Weighted N=2.22M; (3) Design effect=1.56; (4) *estimate significant at $p \leq .01$, otherwise estimates are significant at $p \leq .001$. **SOURCE**: National Center for Education Statistics: High School & Beyond/Sophomore cohort, NCES CD#98-135.

Parenthood, a status that applies to men as well as women, is a frequently overlooked variable in educational histories, or, at best, is checked off as an aside even when it shows a strong relationship to access (as in Akerhielm, Berger, Hooker, and Wise, 1998). The earlier one has children of one's own, the higher the risk of not completing one's education (Horn, 1996; Horn, 1997). This sounds like an easy issue to isolate, but one learns quickly to be careful with student responses on questionnaires. For example, we have people in the HS&B/So who told us in 1984 that they had children, but when asked in 1986 whether they "ever had children," answered in the negative. We have 19 year-olds who graduated from high school in 1982 and told us that their children were born in years such as 1949, 1955, and 1968. The dichotomous variable developed and used in this analysis, "Children," treats all contradictory and (to put it gently) out-of-scope responses as "no children."

Summary

Previous research on the determinants of degree completion has been wanting on pre-collegiate measures and uncritically accepting of stock variables such as aspirations that, on closer examination, require reconstruction. Disaggregating SES into its component pieces may be inviting, but, as the case of parental level of education illustrates, is hazardous: the whole is stronger and more consistent than its parts. At the same time, we ignore the most basic of events in life-course histories, for example, having children in late adolescence or early adulthood, at our peril.

With these issues behind us, we can now bring the college transcripts onto the table of evidence, for they will enable us to grasp the activities of individuals moving through a series of learning environments in late adolescence and early adulthood. As the next portion of our exposition should make amply clear, this mobility has taken on dimensions that render traditional inquiries on the paths to degree completion—how should one say?—quaint.

-III-
The Age of Multi-Institutional Attendance

Beneath the cross-sectional portraits of U.S. higher education that appear annually in *The Digest of Education Statistics* and analogous volumes lie seething movements of students. To break through the pasteboard masks of these portraits we would do well to mark changing attendance patterns over the period, 1972-1995. Three NCES longitudinal studies are the sources for this story, though the HS&B/So is its core. The story advances on other accounts of attendance patterns (McCormick, 1997; Berkner, Cuccaro-Alamin, and McCormick, 1996; Hearn, 1992; Carroll, 1989; Peng, 1977) by distinguishing between "classic" modes of transfer and multi-institutional portfolios, and by combining these variations with temporal factors, including a different way of thinking about part-time status. The increasing complexity of attendance patterns is one of the most significant developments in higher education of our time, one that poses grave challenges to system-wide planning, quality assurance, and student advisement.

The Postsecondary Shopping Mall

The core hypothesis: the long tradition of institutional effects research in higher education is outmoded. The engines of its demise lie in student behavior. The changes in attendance patterns, which can be tracked only in national longitudinal studies, are part of the larger currents of a wealthy open market that produces dozens of specialty niches in every sub-sector. As a society, we have become more consumerist and less attached to organizations and institutions with which we "do business." By consistently selling itself in terms of how much more money students will earn in their lifetimes as a consequence of attendance, higher education has come to reflect other types of markets and marketplaces. As colleges continue to create new specialty majors, dividing academic space times-on-times, they have inevitably drawn a panoply of rival providers (Marchese, 1998). Convenience, of which location is a reflection, has become the governing filter of choice[35], and convenience applies not only to place, but also to time, subject, and price. It is thus not surprising to find students filling their undergraduate portfolios with courses and credentials from a variety of sources, much as we fill our shopping bags at the local mall.

One need not recite the mass of student history studies, starting with Feldman and Newcomb (1969) and Astin (1977, 1993), the integrity of which is predicated on students entering and attending only one institution. The major theoretical models of retention/attrition (e.g. Bean, 1980, 1982, 1983; Tinto, 1975, 1982, 1993) are based on this premise. After all, questions about academic and social growth, let alone those about retention and degree completion, are tainted by second and third institutions. When students disappear from an institution's radar screen, they are assumed to be drop-outs---unless they return. Until recently, we've never known where the stopouts who returned to their first institution of attendance have been during the stopout period. The drop-outs are entered as half of a dichotomous outcome variable in the standard multivariate analyses where institutional effects variables are all drawn from the same school or schools with similar characteristics (e.g. Carnegie class, size, urbanicity) and where demographic effects can be isolated (Astin, Tsui, and Avalos, 1996).

Even transfer from 2-year to 4-year colleges is usually excluded from these statistical models, an unfortunate phenomenon in light of the bachelor's degree attainment of transfer students (Lee, Mackie-Lewis, and Marks, 1993). There are exceptions to this exclusionary practice (Nora, 1987; Lee and Frank, 1990), but institutional variables seem to play less of a role in stories of this "classic" mode of transfer.

Recognition of Changing Attendance Patterns

Analyses based on discrete events in students' lives begin to take account of the problems with traditional institutional effects research. Stop-out and transfer become part of the portrait (Guerin, 1997; Carroll, 1989), and reentry is acknowledged as a critical chapter in students' postsecondary histories (Smart and Pascarella, 1987; Spanard, 1990). Using the Beginning Postsecondary Students (BPS) longitudinal study of 1989-1994, McCormick (1997) takes Carroll's "persistence track" array one step further with an explication of dozens of transfer and multi-institutional sequences by adding three levels of institution (4-year, 2-year, and less-than-two-year) and level and control (public, private not-for-profit, and private for-profit) of the first institution of attendance. In this presentation, there are nine sets of "transfer origin and destination," and these can be subset by seven combinations of level and control.

All of these recent analyses reflect empirical realities. Something is going on. To what extent does that "something" require that we change the way we do research on institutional effects? To what extent does that "something" alter the process of state planning for the provision of higher education? What happens to the standard multivariate model of persistence/completion when different attendance behaviors and different constructions of stock attendance behaviors (e.g. full-time/part-time) are entered? Let us first introduce the variables in the analyses, with descriptive tables that provide some hints of how they might play out in regression equations where bachelor's degree completion is the dependent variable.

Attendance Patterns and a New Universe of Variables

Table 18 documents the changes in the growth of multi-institutional attendance over the past quarter century, The definition of "attendance" is important to table 18. Without completely reconstructing the NLS-72 data base, its definitions were shaped as closely as possible to those used for the HS&B/So. The High School & Beyond variable for number of undergraduate transcripts requested was based on a hand-and-eye examination of the student's consolidated record. Graduate school transcripts were flagged and placed outside the basic calculation. At the same time, we added any institution the student did *not* mention in his/her interview but which was referenced on a transcript from another institution. Transcripts covering only summer school attendance were counted, but only when more than 6 credits in more than two courses were earned. Entries documenting study abroad[36] and cases of transcripts requested from educational institutions in other countries (none were ever received) were also counted as second, third, or fourth schools (only 2.8 percent of all HS&B/So college students were affected). For the HS&B/So, then, we had a fairly strict accounting of attendance. Even if some of that second or third school attendance was incidental–for example, 7 to 10 credits–it was not fragmentary.

I did not anticipate the emergence of a multi-institutional attendance theme when editing the NLS-72 postsecondary student records a decade ago. So the current accounting for the NLS-72 was wholly algorithmic, and involved but a minor (and obvious) adjustment: for students who had earned a B.A., transcripts that showed nothing but post-B.A. course work were not counted, as were out-of-scope responses from schools where transcripts were requested but not received. At the same time, by default algorithm, transcripts that included nothing but summer school work were included. The rate of multi-institution attendance is thus slightly overstated for the NLS-72.

There are some shaggy ends in this process, but the samples are robust enough so that the trends evident in table 18 command a reasonable degree of confidence. Table 18 presents data for two groups: all non-incidental students and bachelor's degree recipients. Bachelor's degree

Table 18.—Change in the percent of students attending one and more undergraduate schools: "Basic Participation Populations"* v. Bachelor's Degree Recipients

	Number of Undergraduate Schools		
	ONE	TWO	>TWO
Basic Participants			
NLS-72 (1972-1984)	59.5%	30.6%	9.9%
HS&B/So (1982-1993)	46.8	33.7	19.5
Change	**-12.7**	**+3.1**	**+9.6**
Bachelor's Recipients			
NLS-72 (1972-1984)	50.4	36.4	13.2
HS&B/So (1982-1993)	41.7	36.7	21.6
Change	**-8.7**	**+0.3**	**+8.4**

Effect size for Basic Participants = .36
Effect size for Bachelor's Recipients = .24

* Basic participation = earned more than 10 undergraduate credits.

NOTES: (1) All within-cohort differences are significant at $p \leq .05$. (2) The effect size is used for measuring differences in means across cohorts. One takes an unweighted sample N, mean, and standard deviation for each cohort group. Then one creates a "pooled standard deviation" by the formula: $(N^{1}(SD^{1}) + (N^{2}(SD^{2}) / (N1 + N2)$. Lastly, one determines the difference in the means and divides that difference by the pooled standard deviation. The resulting effect size functions like a Standard Deviation Unit or z-score, that is, it tells us whether the change in the mean is truly significant. In this case, the effect size for basic participants is moderate and that for bachelor's degree recipients small. For guidance on interpreting SDU-type changes, see Bowen, H.R., *Investment in Learning* (San Francisco: Jossey Bass), 1973, p. 103. **SOURCES**: National Center for Education Statistics: National Longitudinal Study of the High School Class of 1972, and High School & Beyond/Sophomore cohort.

recipients include community college transfers, therefore will evidence a higher degree of multi-institutional attendance than the larger universe of basic participants in higher education. The rate of multi-institutional attendance increased for both groups between the 1970s and 1980s (the effect size clearly shows the increases to be significant, though more for the basic participants than the bachelor's degree recipients), but it is notable that the major shift was not from one to two schools but from one to three or more.

These trends continue. The more recent Beginning Postsecondary Students longitudinal study of 1989-1994 provides evidence of momentum toward even higher rates of multi-institutional attendance. This is obviously a shorter-term study than either of the age-cohort longitudinal studies used in table 18, and includes a more diverse population in terms of age. Since older beginning students are less likely to transfer (see McCormick, 1997, p. 46), table 19 excludes them so as to render the population more analogous to those of the age cohort studies. The table also drops the "basic participation" criterion applied to transcript-based longitudinal studies so as to render them more comparable with BPS. In other words, anyone who ever walked through the door of a postsecondary institution and generated a record--even if they dropped out permanently within a week--is included in table 19. With these criteria, by the fifth year following initial entry, the proportion of BPS90 students who attended more than one institution was already four percentage points higher than the 11-year history of the High School & Beyond/Sophomores.

Ostensible v. "Referent" First Institution

Table 19 begins to mix in institutional type as a factor in the analysis. We want to know whether there are any significant differences or trends in the relationship between the first institution of attendance and the total number of institutions attended. In this respect, there are some bumps in the road across table 19. First, some institutions changed status over the 22-year period covered by the three cohorts. Secondly, the HS&B/So allows one to distinguish between the ostensible first institution of attendance and something we might call the "referent" first institution. The "referent" institution is a retrospective judgment with decision rules. It was determined by hand-and-eye examination of students' records by two readers who could draw on other information from the data set to answer the question: what was the first institution at which this student really "made a 'go' of it?" The determination thus excludes college course taking while the student was still enrolled in high school, enrollment during the summer term following high school graduation at an institution other than the school the student entered in the fall term, and false starts. A "false start" would be indicated, for example, by an initial one-term enrollment at a state college with the student attempting 15 credits and withdrawing from 12 of them, and nothing else on the record until two years later, when the student enrolled in a hospital school for radiologic technology and completed a certificate program. The referent first institution in this case is the hospital school. The "referent" first institution is missing for three percent of the students, principally because their records were incomplete, and these students are excluded from correlations and regression models involving attendance pattern variables.

Table 19.—Percent of students attending more than one institution of postsecondary education, by type of "referent" first institution attended

	NLS-72 (1972-84)	HS&B/So (1982-93)	BPS/90 (1989-94)
First Institution			
4-Year College	38.8*	52.3*	50.1
2-Year College	36.4*	46.5	56.5
Other	7.9	52.5*	30.9
ALL	**35.6**	**48.3**	**51.8**
Sector Share of All Cohort Students:			
4-Year College	54.7	50.5	55.5
2-Year College	37.8	39.7	40.1
Other	7.5	9.7	4.4

NOTES: (1) The universe for the BPS/90 distribution consists of "typical" college-age (17-23) students only, so as to render it more comparable to the age-cohort longitudinal studies. (2) The universes for both the NLS-72 and HS&B/So consist of all students with postsecondary records. (3) Definition of first institution of attendance for the HS&B/So differs from that of the other studies. See text. (4) All within-cohort differences are statistically significant at $p \leq .05$ *except* those column pairs indicated by asterisks. SOURCES: National Center for Education Statistics: (1) National Longitudinal Study of the High School Class of 1972; (2) High School & Beyond/Sophomore Cohort, NCES CD #98-135; (3) Beginning Postsecondary Students, 1989-1994, Data Analysis System.

Approximately 10 percent of the HS&B/So students who entered postsecondary education are affected by the "referent" first institution distinction, that is, there is a difference between their ostensible date/place of entry and their "true" date/place. But these effects are not evenly distributed, and there is a clear tone to the distribution: overrepresented are students who entered selective and doctoral degree-granting institutions, those who came from the highest

SES quintile, and those who eventually earned associates and/or bachelor's degrees. These students were more likely than others to earn college credits while they were still in high school and/or during the summer immediately following high school graduation. If we had not created the "referent" variable, estimates for these students would have been distorted, and, because the size of the group is not insignificant, these estimates would affect multivariate analysis. For example, the majority of these students would show either a community college or a non-selective 4-year college for the ostensible first institution of attendance. In a multivariate analysis, controlling for other attendance pattern variables, that would result in underestimating the role of the selectivity and type of the first institution.

Institutional Combinations

More than one college means not only schools of different types, but also in different states (and countries), in alternating and simultaneous patterns. Table 20 presents a portrait of those patterns derived from the HS&B/So transcript files. It is immediately obvious that traditional factors involved in the analyses of academic and social integration (frequency of contacts with faculty, staff concern for student development, peer relations, and so forth) take on a very different coloration among students who attend more than one school, particularly if they did not return to the first institution of attendance--something we've always known about community college transfers who spent any appreciable time at the community college. The universe of variables involving key institutional effects contracts as the number of institutions attended rises, even when students return to the first institution of attendance (as did 61 percent of those who attended two schools, and 48 percent of those who attended three or more).

Drawing on the Beginning Postsecondary Students study of 1989-1994, McCormick (1997) employs a different scheme but with a similar tone and results. In that five-year span (as opposed to the eleven years of the HS&B/So), the 45 percent of students who had enrolled in more than one institution as undergraduates displayed most of the inter-sectoral and intra-sectoral patterns shown in the "institutional combinations" section of table 20. The considerable mobility evidenced by the HS&B/So student population, then, was not a temporary blip on the radar screen of higher education.

Table 20 uses nine categories of institutional combinations, six of which describe inter-sectoral movement, and caps the final date for calculation at that of the bachelor's degree (if earned) or (if the bachelor's was not earned) at the last term date of undergraduate attendance. Temporal considerations enter in the construction of three of the inter-sectoral categories in order to clarify both enduring and emerging policy concerns. The first involves the notion of "reverse transfer" under which the referent first institution of attendance is a 4-year college and the last institution is a 2-year college, with no bachelor's degree earned in between. Students in the HS&B/So who attended a community college after earning the bachelor's degree sport a separate flag on their records. Only 3.3 percent of the bachelor's degree recipients in this cohort (weighted N=26,751) carry this flag[37].

Table 20.—A portrait of multi-institutional attendance of the HS&B/So, 1982-1993: percent of students attending more than one college, by institutional characteristics and other attendance variables

	Number of Undergraduate Institutions				
	One	Two	Three+	% of All	% Earning Bachelor's
All who earned >10 credits	47	35	18	100%	45
Type of Referent First Institution					
Doctoral	44	36	20	23	74
Comprehensive	49	32	19	22	59
Liberal Arts	43	38	19*	6	73
Community College	46	34	20	34	20
Other	63	24	13*	15	11
Selectivity of First Institution					
Highly Selective	51	31	18*	3	92
Selective	44	36	20	9	85
Non-Selective	44	36	20	52	65
Open Door	56	28	16	31	7
Not Rated	75	18*	7*	5	2
Institutional Combinations					
4-year only	56	31	13	46	75
4-year to 2-year	--	71	29	4	--
2-year to 4-year	--	63	37	10	66
Alternating/Simultaneous	--	30	70	5	52
4-year and other	--	78	22*	3	17*
2-year only	75	18	7*	22	--
2-year and other	--	77	23	3	--
Other only	90	10*	--	6	--
2-year, 4-year & other	--	--	100	2	14*
Timing of Entry					
No Delay	45	35	20	81	51
7-18 Month Delay	52	29	19	9	20
>18 Month Delay	70	22	9*	10	10

Notes: (1) The universe consists of all students who earned more than 10 credits. Weighted N=1.87M. (2) "Other" institutions consist principally of non-degree granting vocational schools, and these are "not rated" in terms of selectivity. (3) Institutional type based on the Carnegie classification system in effect in 1987. (4)* Low-N cells. Estimates are not significant. **SOURCE**: National Center for Education Statistics: High School & Beyond/Sophomore cohort, NCES CD#98-135.

The second of these inter-sectoral values is that of traditional 2-year to 4-year transfer. This value was assigned only on the basis of sequence, without consideration of the number of credits earned at each institution or what proportion of those credits were accepted in transfer (something often not indicated on the record of the receiving institution). The "transfer" variable used in the multivariate analyses in Part IVof this study is *not* derived from this simple configuration, rather from a construction involving credit thresholds developed for analyses of the NLS-72 (Adelman, 1994). This construction requires the student to start at a community college, earn more than 10 credits from the community college, and subsequently attend a 4-year college and earn more than 10 credits from that institution.

Some students assigned the Alternating/Simultaneous enrollment value turn out to be "classic transfer" students. The "alternating" portion of this category involves 2-year and 4-year institutions only, isolating the fact that not all transfers from 2-year to 4-year institutions involve one event. The other portion of the category includes *any* case of simultaneous enrollment (determined by overlapping term dates on transcripts from two institutions). Students in the Alternating/Simultaneous category exhibit a very high proportion of attendance at *at least three* institutions. Alternating attendance is also possible among institutions of the same type, and if we put these cases together with those of the inter-sectoral Alternating/Simultaneous category, we describe the oscillations of 16 percent of the HS&B/So postsecondary universe, and nearly 18 percent of those who earned bachelor's degrees. These are not insignificant portions.

The Issue of Return

One of the key questions involving any case of multi-institutional attendance is whether the student *returned* to the referent first institution. As we'll see in the multivariate analyses, the variable developed with this notion in mind is stronger than a mere count of undergraduate institutions. If one attends two or more schools and does *not* return to the first, one has engaged in a very different kind of movement than a stroll around the neighborhood.

The variable that captures this phenomenon was constructed by defining the date of the last term of undergraduate enrollment, and then determining whether the institution at that last date was the same as the referent first institution of attendance. There are some minor problems with any algorithm such as this, but the results reinforce common sense. For example, the more schools attended, the less likely the "return rate," and those who do *not* start in 4-year institutions are less likely to return to their first institution (this group includes the classic transfer students who, by definition, do not return to their first institution).

It might be helpful to sort the common-sense from the counter-intuitive with a correlation matrix confined to the major attendance pattern variables for students who attended more than one institution. Table 21 offers this matrix. Five dichotomous measures of place and two of time are included. The dependent variable is ANY 4-YR, that is, whether the student ever attended a 4-year college. The other variables in the matrix are:

NUMSCHL The dichotomy splits students who attended only 2 schools from those who attended 3 or more.

FIRST4 The student's referent first institution was a 4-year college;

INSYS The student's multi-institutional attendance was confined to the system of 2-year and 4-year colleges. The student never stepped outside the system.

NODELAY The student's referent date of entry to postsecondary education occurred within 10 months of high school graduation.

NOSTOP Continuous enrollment. The student never "stopped out" for more than two semesters or three quarters (see p. 53 below).

Table 21.—Correlations of major attendance pattern variables for students who attended more than one postsecondary institution: High School & Beyond/ Sophomores, 1982-1993

	No Return	Numschl	First 4-Yr	Any 4-Yr	In System	No Delay	No Stop
No Return	1.000	-0.146	-0.072	0.032	-0.107	0.030	-0.099
p		0.001	0.02	----	0.01	----	0.01
Numschl		1.000	-0.004	-0.131	-0.038	-0.030	0.124
p			----	0.001	----	----	0.001
First 4			1.000	0.541	0.212	0.322	0.279
p				0.001	0.001	0.001	0.001
Any 4-yr				1.000	0.276	0.342	0.338
p					0.001	0.001	0.001
In System					1.000	0.129	0.143
p						0.001	0.001
No Delay						1.000	0.321
p							0.001

<u>NOTES</u>: (1) Universe consists of all students who earned more than 10 undergraduate credits and attended more than one postsecondary institution. (2) Weighted N=1.069M. (3) Design effect=1.52. <u>SOURCE</u>: National Center for Education Statistics: High School & Beyond/Sophomore cohort, NCES CD# 98-135.

What does table 21 begin to tell us? The correlations are not very high, but most are statistically significant. Some of the messages indicate that cross-currents are at work. The correlation coefficients suggest that:

● The temporal dimensions of attendance (delay and continuity) are more strongly connected to the principal features of 4-year college participation than are those reflecting number of schools or sequence of attendance. In multivariate analyses confined to students who attended 4-year colleges, we would thus expect the temporal variables to contribute more to a model explaining degree completion.

● Students who attend 4-year colleges are less likely to step outside the secular higher education system. In multivariate analyses confined to 4-year college students, we would thus expect whatever variable we create to represent this phenomenon to fall out of model equations explaining either the variance (in linear equations) or likelihood (in logistic models) of bachelor's degree completion at an early iteration.

● The fewer schools attended the more likely continuous enrollment (a common-sense hypothesis), and the less likely a 4-year college is part of the attendance pattern. If we confined a multivariate analysis to students who had attended a 4-year college at any time, the first of these dimensions would be a strong contributor to a model of degree completion.

Mobility: Sweeping State Borders Away

Most of the variables in table 21 are those of place--at least in a generalized sense: number of schools, sector, type. One prominent aspect of place that is not covered by this configuration is geographical, and because states and regional bodies such as accrediting associations and interstate compacts (Western Interstate Commission on Higher Education, the Southern Regional Education Board, the New England Board of Higher Education) play key roles in planning or advising for the provision of higher education, a portrait of *multi-state* attendance may be helpful. After all, a major problem in discussions and arguments over graduation rates is that even the best tracking systems are confined within state borders. Given the extent of multi-institutional attendance, *system* graduation rates are very different from institutional graduation rates. Even statewide rates--for students who began their higher education careers in a given state--do not capture the movement of students across state lines. Table 22 sets forth the inter-state dimensions of mobility in the careers of the HS&B/So.

One category of analysis in table 22, "aggregate pattern categories," requires some explication. The "inter-sectoral" pattern describes any non-recursive (one-way) change of institution that crosses a border from the 4-year college sector to the 2-year college sector (and vice versa), or the "non-secular" sector (and vice versa). Students evidencing a classical community college to 4-year college transfer are included here. The "intra-sectoral" pattern describes non-recursive multi-institutional attendance *within* a given sector. An "alternating" attendance pattern is one in which the student is oscillating between two or more institutions, no matter what type of institutions they may be. The in*tra*-sectoral pattern is slightly more likely to involve at least two states than the other two patterns. It will not surprise anyone that the vast majority (82 percent) of students in the intra-sectoral pattern who cross state lines are moving from one 4-year college to another.

Table 22.—Percent of multi-institutional attendance students who attended college in one or more states, by other attendance pattern variables, High School & Beyond/Sophomore cohort, 1982-1993

	Number of States		
	One	Two	Three or More
ALL:	61.7%	32.8%	5.5%
Total Number of Schools			
Two	68.6	31.4	----
Three or More	48.7	35.4	15.8
Aggregate Pattern Categories			
Alternating	61.1*	29.9*	10.0
Inter-Sectoral	66.4*	28.1*	5.5*
Intra-Sectoral	57.5	38.4	4.1*
First Institution			
4-Year	54.9	39.0	6.1*
Not 4-Year	69.0	26.2	4.8*
Selected Institutional Combinations			
4-Year Only	51.4	42.7	5.9
4 to 2-yr transfer	66.7#	25.1*	8.2
2 to 4-yr transfer	74.2*	22.2	3.6
2-Year Only	73.8*	24.7*	1.6
Other	69.0#	31.0#	---
4-Year/2-Yr/Other	46.5	29.2#	24.4

NOTES: (1) Differences in all possible row-pairs are significant at $p \leq .05$.
(2) All column pair comparisons are significant at $p \leq .05$ *except* those indicated by * or #. (3) The universe consists of all HS&B/So students who earned more than 10 credits and attended more than one postsecondary institution as undergraduates.
Weighted N=1.069M. **SOURCE**: National Center for Education Statistics: High School & Beyond/Sophomore cohort, NCES CD#98-135.

Table 22 confirms other basic observations: the more schools one attends and the more a 4-year college is involved, the more likely one is to cross state lines in the process. The swirl of multi-institutional attendance sweeps all borders away. Furthermore, for students who attended more than one college (including at least one 4-year college), the bachelor's degree attainment rate for those who crossed state lines was higher (62 percent) than for those who stayed within state borders (55.4 percent). These issues should be important to state planners who are currently drawing on a variety of demographic scenarios to predict likely future enrollments in higher education. The most noted of these scenarios, a joint study of the Western Interstate Commission on Higher Education (WICHE) and the College Board (1998), focuses on the supply of traditional-age high school graduates. Even when they predict the high school graduating population in a state, however, and even if they predict the proportion of those students who will continue on to college, none of the current modeling exercises includes post-matriculation behavior such as multi-institutional attendance patterns and inter-state enrollment commerce. The High School & Beyond/So histories, including these features, on the other hand, have been shown to be helpful with *national* projections (Adelman, 1999). However useful these are as background tapestries, the HS&B/So was not designed for—and cannot be applied to—state based analyses.

Further Factors of Space and Time

The literature on persistence and degree completion places considerable emphasis on variables describing timing of college entry and the type of institution first attended. A number of conventional wisdoms have grown into folklore from these research lines: starting in a 4-year college produces a higher bachelor's degree completion rate than starting in a 2-year college; the more selective the first institution of attendance, the higher the completion rate; the greater the delay between high school graduation and college entry, the lower the completion rate. Let us see how this folklore plays out in an analysis that refines both of these issues with the notion of "referent" first institution and—hence—"referent" first time.

First Institution: Sector and Selectivity

For purposes of basic understanding, the characteristics of first institution can be parsed by sector (4-year, 2-year, other), type (doctoral, comprehensive, liberal arts, community college, specialized, and other), selectivity (highly selective, selective, non-selective, open-door, and not rated), and level/control. I have never found level/control to be as revealing as the other basic characteristics. For example, in the standard taxonomy attached to all the national longitudinal studies data sets, proprietary schools that offer bachelor's degrees (including music conservatories, art schools, and large technical school chains) are *not* included with 4-year institutions. If, as the literature suggests, the 4-year school is the critical first institution of attendance for bachelor's degree completion and we scramble what we mean by a 4-year school, then we distort subsequent analyses and interpretations. A for-profit music conservatory can be highly selective. What is more important in considering the impact of attending such an institution: its for-profit status or the fact that it offers a bachelor's degree and is selective?

Table 23.—Proportion of students earning bachelor's degrees by sector of "referent" first institution of attendance: High School & Beyond/Sophomores, 1982-1993

	Access* Group	Participation* Group
Sector of 1st Institution		
4-Year	63.2 (1.14)	65.9 (1.11)
2-Year:		
All Students	14.8 (0.85)	19.3 (1.09)
Attended 4-Year at Any Time	59.0 (2.35)	60.0 (2.37)
Transfer Students	71.1 (2.33)	71.1 (2.33)
Other	4.4 (0.93)	5.7 (1.22)
4-Year College Sector		
Doctoral	72.0 (1.57)	73.5 (1.11)
Comprehensive	54.7 (1.62)	58.7 (1.66)
Liberal Arts	71.4 (2.59)	73.0 (2.55)
Specialized	42.5 (5.60)	43.7 (5.76)

NOTES: (1) *The "access group" consists of anyone for whom we received a transcript, even if there were no credits on the transcript (Weighted N=2.27M); the "participation group" consists of all students who earned more than 10 undergraduate credits by age 30 (Weighted N=1.94M). (2) Standard errors are in parentheses. (3) Transfer students, by the definition used in this study, earned more than 10 credits from a 2-year college and, subsequently, more than 10 from a 4-year college. **SOURCE**: National Center for Education Statistics: High School & Beyond/Sophomore Cohort restricted file, NCES CD#98-135.

Table 23 offers an account of bachelor's degree completion rates by sector and type of the first institution of attendance. It also hones in on students who started in 2-year colleges in order to highlight some rather dramatic differences within this population. For all students who began their postsecondary careers in a community college and who earned more than 10 credits, about one out of five eventually earned a bachelor's degree. The denominator of this ratio, however, includes a mass of students who never attended a 4-year college and had no intention of doing so. On the other hand, it is rather obvious that students who began in a community college, earned more than 10 credits from the community college, and subsequently attended a 4-year institution, whether in a classic transfer pattern or in an alternating/simultaneous enrollment pattern, eventually earned bachelor's degrees at a rate (71 percent) higher than that for those who started in a 4-year college. "Early transfers," those who jumped ship from the community college to the 4-year institution with 10 or fewer credits, completed bachelor's degrees at a much lower rate--38.4 percent (this group is not shown in table 23). Table 23 also suggests a bi-modal pattern within the 4-year college sector, with clearly higher degree completion rates for those who start in doctoral and liberal arts institutions than those whose referent first institution is a comprehensive or specialized college. Because the proportion of 4-year college students whose first institution of attendance was a liberal arts college is low (12 percent), when we set up a dichotomous variable for institutional type in a multivariate equation, the construction is "doctoral" versus all others.

Table 24 moves on to the parsing of first institution of attendance by generalized selectivity levels for students in the "participation group." The selectivity levels were determined from the Cooperative Institutional Research Project's selectivity cells for first-time freshmen in 1982 (Astin, Hemond, and Richardson, 1982), the year most of the HS&B/So cohort entered higher education. For institutions not in the 1982 CIRP, selectivity was determined from the 1982 edition of Barron's *Profiles of American Colleges*. Given these different sources, a trichotomy

Table 24.—Distribution of students by selectivity of (1) "referent" first institution of attendance and (2) bachelor's degree awarding institution, and completion rates by selectivity of first institution, High School & Beyond/Sophomore cohort, 1982-1993

	Highly Selective	Selective	Non-Selective
At First Entry	4.0%	13.7%	82.3%
At Graduation	5.5	17.1	77.4
BA Completion Rate for First Entrants	93.1*	87.9*	64.7

NOTES: (1) The universe for "first entry" is confined to students whose referent first institution was a 4-year college and who earned more than 10 undergraduate credits. Weighted N=1.03M. The universe for graduation is confined to those for whom a bachelor's degree is documented. Weighted N=935k. (2) All row pair comparisons are significant at $p \leq .05$ *except* those indicated by asterisks. **SOURCE**: National Center for Education Statistics: High School & Beyond/ Sophomore cohort, NCES CD#98-135.

seemed to be the most appropriate way to establish statistically significant borders. The effects of selectivity are enormous, though they may be confounded by the level of academic resources students bring forward from secondary school (a hypothesis to be explored in a multivariate context). These selectivity effects are reflected not only in bachelor's degree completion rates, but also in the proportion of graduates who continue to graduate or professional school, and in undergraduate GPA.[38] With respect to bachelor's degree completion rates, there is no real difference between the "highly selective" and "selective" institutions, thus justifying a dichotomous variable in multivariate analyses (any selectivity versus none).

Continuity of Enrollment

Continuity of enrollment is the first of two temporal variables that deserve further description. The definition of continuity of enrollment depends on the length of the period of measurement. Carroll (1989), Berkner, Cuccaro-Alamin, and McCormick (1996), and others, working with longitudinal studies of five or six year time-frames, define non-continuous enrollment ("stop-out") as a gap of four months or more (excluding summer periods) in spells of attendance. The postsecondary time-frame for the High School & Beyond/Sophomores, however, is 11 years. While the vast majority of the cohort, including those who did not earn degrees, drift away from postsecondary education after 8 years (age 26/27), we do not take the benchmark measure of completion until the history is censored at age 29/30. Under those circumstances, and based on a hand-and-eye reading of the records by two judges, a student's enrollment was judged to be non-continuous if it evidenced a break of (a) two or more expected consecutive semesters, or (b) three or more expected consecutive quarters, or (c) two or more breaks of one expected consecutive semester or two expected consecutive quarters. Put another way, one no-year or two or more part-year enrollment spells are thus the units of analysis behind the judgment of non-continuity.

The continuous enrollment variable in the HS&B/So data base is not dichotomous. Students who were enrolled for less than one year (6.4 percent of all who entered) are not subject to judgment, and were assigned a separate value. And there were many cases (11.4 percent of all students) where continuity could not be determined, usually because records or term dates were missing. In presenting a descriptive account of continuity of enrollment, table 25 serves a dual role. It first maps continuity of enrollment against the aggregated multi-institutional attendance variable for students who earned more than 10 credits; and then provides guidance for constructing a dichotomous version (for convenience, called NOSTOP) for use in multivariate analysis of bachelor's degree completion.

Part A of table 25 provides further confirmation of what we have observed of the positive relationship between 4-year college attendance and continuous enrollment, as well as the negative relationship between inter-sectoral (including alternating/simultaneous) patterns of multi-institutional attendance and continuous enrollment. Part B of table 25 provides a preview of the potential strength of NOSTOP in multivariate analyses: the completion rate for continuously enrolled students is two times that for non-continuously enrolled students.

Table 25.—Proportion of students' enrollment continuity status by attendance pattern and bachelor's degree completion rates, High School & Beyond Sophomore cohort, 1982-1993

	Enrollment Continuity Status			
	Con-tinuous	Non-Contin	Less Than 1 Year	Can't Determine
Part A. Attendance Pattern				
Alternating/Simultaneous, all Sectors	45.8%*	43.8%*	0.0	10.4%
Inter-Sectoral	41.5*	39.4*	1.3	17.8
Intra-Sectoral	67.0	26.1	0.3	6.6
1 Institution Only	77.8	16.5	5.7	---
Students Who Attended a 4-year College At Any Time	70.2	23.2	4.9	1.7
Part B. Completion				
Bachelor's Degree Completion Rate for Students Who Attended a 4-Year at Any Time	77.3	38.5	N.A.	0.2

NOTES: (1) The Part A universe consists of all students who earned more than 10 credits and for whom a value of continuity could be determined. (2) Column comparisons for Part A are significant at $p \leq .05$ *except* for the asterisked pairs. (3) Rows for Part A add to 100.0%. **SOURCE**: National Center for Education Statistics: High School & Beyond/Sophomore Cohort, NCES CD#98-135.

Full-Time, Part-Time, or DWI?

Whether a student is full-time or part-time is a stock variable in virtually all analyses of postsecondary persistence and attainment. To repeat an observation from Part II (pp. 30-31 above): student reports of part-time enrollment are severely lower than aggregate institutional reports; the FT/PT measure is usually based on a snapshot of the first term of attendance, and, most importantly, students change enrollment status during their college careers.

Transcript analyses render the notion of "part-time" even more fragile. Students may start a term with a 15-credit load (full time), but drop or withdraw from six or more credits before the term is over. One never knows precisely when the drop or withdrawal took place, but a student who behaves in this manner moves into part-time status in the course of the term. It would be enormously difficult to estimate the true part-time rate in longitudinal studies, but we can draw some parameters from the transcripts. Table 26 demonstrates what I call the "DWI Index" for the HS&B/So cohort by highest degree attained by age 30. What we see is the percent of students who either dropped, withdrew from (with no penalty), or left as unresolved "incompletes," proportions of the courses for which they registered over the course of their undergraduate careers. The proportions are set in three ranges. At a 20 percent DWI rate, there is a strong likelihood that the student *became* part time at some point. Some 10.6 percent of the students, in fact, evidence a DWI Index in excess of 40 percent, and 95 percent of this group earned no degree.

While Carroll (1989) did not use these data, they support his contention of the "hazards" of part-time enrollment. Whether they are as strong as his other "hazards" remains to be seen. But it will not surprise anyone that of the 20 course categories (out of over 1,000) with the

Table 26.—DWI* Index: percentage of HS&B/So students who dropped, withdrew, or left incomplete proportions of attempted undergraduate courses in three ranges, by highest degree earned, 1982-1993

Proportion of All Attempted Courses that Became DWIs:	$\leq$10%	10-20%	$\geq$20%
For ALL Students:	**62.0**	**14.7**	**23.3**
Highest Degree Earned:			
None	48.1	14.2*	37.7
Associate's	64.0	20.3	15.7
Bachelor's	78.4	14.9*	6.7
Graduate	87.8	9.6	2.6

Notes: (1) All row and column differences significant at $p \leq .05$ except the column pair indicated by asterisks. (2) Universe consists of all HS&B/So students with complete postsecondary records. Weighted N=1.65M. (3) Rows add to 100.0%.
SOURCE: National Center for Education Statistics: High School & Beyond/ Sophomore cohort, NCES CD#98-135. *DWI=Drops, Withdrawals & Incompletes.

highest rates of DWI, eight are remedial courses such as developmental math and remedial reading (Adelman, 1995, p. 269). Students with high DWI indices are carrying other hazardous baggage. Any study of persistence and completion that limits its subjects to full-time students and assumes this variable to be incorruptible (e.g. Chaney and Farris, 1991) risks shaky conclusions. Even McCormick's (1999) transcript-based analysis defines part-time students in terms of credits *attempted* in the first semester, thus overlooking students with a *de facto* part-time status by virtue of DWI.

The volume of DWI (Drops-Withdrawals-Incompletes) phenomena reflected in table 26 above was striking enough[39] to demand a measure, particularly in light of a ratio of Withdrawals to Drops (many of which are by-products of early term scheduling adjustments) of 12:1. No doubt that some of the students with significant DWI activity were part-time to begin with. But others "went part-time." Only 40 percent of the total adjusted sample went through their college careers without a D, W, I, or NCR (No Credit Repeat—included because a repeated course involves the equivalent of one withdrawal) grade. Some 23.2 percent had a DWI Index of .2 or higher. These students reduced their course-taking (and, probably, enrollment-status-qualifying credits) by 20 percent or more. For multivariate analysis, the DWI Index was turned into a dichotomous variable with a threshold of .2. In another stunning case of common sense, its correlation with degree completion is strong--and negative.

Summary: Parts I-III

We have now walked through the major variables of attendance patterns: those that deal with place and those that deal with time. Based on student transcript records, the construction has been inductive and empirical. I have illustrated relations among these variables, though not exhaustively; suggested just how complex these swirling patterns of student behavior have become; and have provided some clues as to how they might play out in explaining bachelor's degree attainment. Given the extent of multi-institutional attendance, I have declined to attach to each institution such characteristics as size, ethnic composition, special mission, curriculum offerings, proportion of students in residence, and other features that appear in the literature on college persistence. If attendance patterns involving more than one institution turn out to play a significant role in explaining degree attainment, then and only then would we be justified in turning to the task of accounting for combinations of institutional characteristics. But the student is the subject, and as we turn to the multivariate analysis, the student's critical life events and judgments move to center stage.

IV. Does It Make Any Difference? Common Sense and Multivariate Analysis

Do the complex patterns of attendance described above have any impact on bachelor's degree completion, and, if so, which variables survive in regression analyses and how much explanatory power do they contribute to a model that also includes pre-college characteristics, measures of early postsecondary performance, modes of financial aid, and satisfaction with major aspects of higher education?

We have now arrived at multivariate analysis. In this section of the study, there will be six sequential ordinary least squares (OLS) regression models:

- Background: high school performance, aspirations, demography (6 variables)
- Financial Aid Modes (3 variables)
- Attendance Patterns (9 variables)
- 1st True Year Performance (3 variables)
- Continuing Performance Effects (3 variables)
- Satisfaction Indicators (5 variables)

As each configuration of variables is entered, the strong stories rise to the top in terms of contribution to the model and statistical significance, while the weak ones fade. At the end of this process, out of 29 independent variables introduced, only 11 remain. The story told by those 11 variables provides very strong guidance for improving degree completion rates for all populations, and particularly for minority and low-SES students.

A logistic regression version of the first five equations is then presented. Logistic regression is the statistician's method of choice when the outcome is a dichotomous variable such as did/did not earn bachelor's degree. Logistic regression expresses itself in a different way than does OLS. Its objective is to identify the "maximum likelihood" of a relationship, the "probability of observing the *conditions* [ital mine] of success" (Cabrera, 1994, p. 227). The principal metric in which it tells its story is an odds ratio—the way in which HIGHMATH was presented above (see p. 17). In some ways, a logistic regression is like an epiphany: its results make a dramatic statement, the parameters of which are sometimes unexpected. OLS, on the other hand, seeks to minimize the difference between predicted and observed probabilities, that is, between the rational and the empirical. The result is hardly analogous to an epiphany: it is slow and unfolding (Pedhazur, 1982), and expresses its fundamental conclusion in a metric accessible to the general reader: the "percent" of variance accounted for by the model (the R^2). Logistic regression has an analogous metric of conclusion, the G^2, but it is less accessible.

Who is in the Universe?

The population to be considered in the mutilvariate equations consists only of those students in the HS&B/So who attended a 4-year college at any time, even though they might also have attended other types of schools. In a longitudinal study that continues to age 30, the mark of someone who intends to earn a bachelor's degree is actual attendance at a bachelor's degree-granting institution, not a statement. These students may have attended other types of schools, but if the dependent variable is bachelor's degree completion, we distort our understanding of what makes a difference if we include people who never really tried.

But what about those students who expected to earn a bachelor's degree, but never attended a 4-year college by age 30? Aren't they students for whom family income and SES play roles much stronger than ACRES or any of its components? Should they be included in the analysis of degree completion, and, if not, what impact will the exclusion have on both our assessment of the validity of the ACRES variable and the analysis of what makes a different in degree attainment. The best way to confront these questions is by comparing two groups of students (1) whose referent first institution of attendance was a community college, (2) who earned more than 10 credits from the community college (that is, they were not incidental students), and (3) who expected to earn a bachelor's degree. One group ultimately attended a 4-year college; the other did not. They are of roughly equal size (a weighted N of 84k for those who did not attend a 4-year college; a weighted N of 96k for those who did). What do they look like? Table 27 sets forth some basic parameters.

Table 27.–4-year or no 4-year? A comparative portrait of students whose referent first institution was a community college, who expected to earn a bachelor's degree, and who earned more than 10 college credits, High School &Beyond/ Sophomore Cohort, 1982-1993

Percent by SES Quintile	High	2nd	3rd	4th	Low
No 4-Yr	21	27	23*	22	7*
4-Year	29	36	22*	9	*
Percent by Academic Resources Quintile					
No 4-Yr	11	27	37	19	6*
4-Year	32	32	27	7	3*
Percent by Total Undergraduate Credits	11-29	30-59	60-89	90+	
No 4-Yr	32	37	22	9	
4-Year	3	7	8	82	
Percent by Total Credits from CommColl					
No 4-Yr	33	36*	21	10*	
4-Year	16	30*	46	8*	

NOTES: (1) All column comparisons are statistically significant at $p \leq .05$ except those indicated by asterisks. (2) Bachelor's degree expectation as indicated in 12th grade. (3) Rows may not add to 100% due to rounding.
SOURCE: National Center for Education Statistics: High School & Beyond/Sophomore Cohort, NCES CD#98-135.

A more dramatic way of illustrating what makes a difference for these two groups is through a logistic regression with attending a 4-year college as the dependent variable.

Table 28.–Among students who began in a community college, aspired to a bachelor's degree, and earned more than 10 college credits, the relative strength of family and high school background in relation to ultimate attendance at a 4-year college: a logistic analysis

	(Beta) Estimate	s.e.	*t*	p	Odds Ratio
Intercept	-3.2456	.604	3.61		
Family Income	-0.0934	.078	0.80	—	0.91
SES	0.4361	.120	2.43	.05	1.55
Academic Resources	0.6031	.120	3.37	.02	1.83

NOTES: (1) Weighted N=181k; (2) standard errors adjusted for design effect; (3) Design effect=1.49. SOURCE: National Center for Education Statistics: High School & Beyond/Sophomore cohort, NCES CD #98-135.

People, including those who aspire to bachelor's degrees, attend community colleges for many reasons. Within the groups at issue here, family income does not play a role in whether they attend 4-year colleges as well. The odds ratio for family income is very close to 1.0 (which indicates no influence) and the parameter estimate does not meet the criterion for statistical significance at all. SES, which transcends income, is significant; but academic resources seem more significant. What do we make of this?

In both the linear and logistic model series below, SES exerts a modest but declining influence on bachelor's degree attainment as students move into postsecondary education and through their first year. The group of students who started in a community college, expected to earn a bachelor's degree, but never attended a 4-year college by age 30 are much weaker in academic resources than their peers who eventually did attend a 4-year college (table 27). They also exhibit a lower SES distribution, though not dramatically so. If we include them in the multivariate analysis, we can speculate that SES would exert the same modest but declining influence on degree completion but at a slightly higher level. For example, in the first stage of the logistic regression series on pp. 80-81 below, SES might carry an odds ratio of 1.26 to 1 instead of 1.22 to 1. The difference is not a compelling reason for including this group in an analysis designed to help us select the tools we will need to use in both maintaining and bringing greater equity to bachelor's degree completion rates.

So, the universe will be limited to those who attended a 4-year college at any time. This sounds simple enough. But in programming the database, there are two potential definitions of the universe of students "who attended a 4-year college at any time." By one definition, we admit only those students for whom we actually received at least one transcript from a 4-year college.

By the second definition, we also include cases where one or more transcripts from 4-year colleges were requested but none were received. For the universe created by the second definition, we have all variables that were self-reported in the surveys, for example, all the financial aid modes, and even values for attendance pattern variables derived simply from the number and nature of transcripts requested. However, the small expansion group begins to drop out of the analysis when we reach the Attendance Pattern model because, without transcripts, it is impossible to determine a value for any variable based on dates, for example, continuity of enrollment and delay of entry, let alone those based on credits, grades, and course-taking[40].

The background model—before introduction of attendance, performance and other college experience variables—is offered in table 29. It consists of three standard demographic constructs: SES (in quintiles), RACE (dichotomous Black/Latino/AmerInd v. White/Asian), and SEX (male=1). It also includes a dummy variable, children, marking whether the individual became a parent at anytime up to 1986 (age 22/23)[41], the composite ACRES (academic resources) to indicate the quality of the student's performance (curriculum, class rank, and test scores) in secondary school, and the sharpened construct of educational "anticipations," as explained on pp. 33-34 above. The dependent variable is bachelor's degree completion by age 30 in 1993.

What does this basic background model say, no matter which universe one uses? In the absence of any other information: (1) the six independent variables in the equation explain between 21 percent (Part A) and 23 percent (Part B) of the variance in long-term bachelor's degree completion among students in the HS&B/So who attended at least one 4-year college at any time up to age 30; (2) of the six independent variables in the equation, that which carries the student's high school background, ACRES, contributes most to the explanation; and (3) of the remaining five independent variables, only the fact of becoming a parent prior to age 22[42] and parents' socioeconomic status contribute anything else of significance to the explanation.

This is the most basic of common sense matters. If the universe had not been limited to students who attended 4-year colleges at some time, ACRES would have contributed slightly more to the explanation of variance in degree completion. When the outcome is degree completion, who you are is less important than the amount and quality of the time you invest in activities that move you toward that goal.

While race and sex fall out of the model very early, no matter which universe of 4-year college students we use, the "anticipations" variable stays in the equation, however marginally. It stays in the equation only because it was set up as a dichotomy with the positive value confined to "bachelor's consistent" expectations, as appropriate to students who attended 4-year colleges. It is for this reason, along with the fact that the next set of variables to be introduced (those reflecting financial aid and student employment while enrolled) do not rely on transcript data that the universe of Part B will be carried forward.

Table 29.—Background Model: The relationship of pre-college and family variables to bachelor's degree completion among 4-year college students in the HS&B/So, 1982-1993

Part A: Restricted Universe of 4-Year College Attendees

Universe: All students for whom a transcript from a 4-year college was received, who received a high school diploma or equivalent prior to 1988, and who evidenced positive values for all variables in the model. N=4,765. Weighted N=1.131M. Simple s.e.=.702; Taylor series s.e.=1.083; Design effect=1.54.

Variable	Parameter Estimate	Adj. s.e.	t	p	Contribution to R^2
INTERCEPT	1.782346	.1159	9.98		
ACRES	0.133301	.0094	9.21	.001	.1651
Children	-0.340264	.0433	5.10	.01	.0282
SES Quintile	0.036144	.0083	2.83	.05	.0122
Anticipations	0.071447	.0218	2.13	.10	.0039
Race	-0.073809	.0300	1.60	---*	.0026
Sex	-0.040801	.0193	1.37	---*	.0017
*Dropped from model.				R-Sq.	.2137
			Adj.	R-Sq.	.2127

Part B: Unrestricted Universe of 4-Year College Attendees

Universe: All students for whom the evidence confirms 4-year college attendance at any time (whether transcripts were received or not), who received a high school diploma or equivalent prior to 1988, and who evidenced positive values for all variables in the model. N=4,943. Weighted N=1.179M. Simple s.e.=.697; Taylor series s.e.=1.082; Design effect=1.55.

Variable	Parameter Estimate	Adj. s.e.	t	p	Contribution to R^2
INTERCEPT	1.722264	.1112	9.99		
ACRES	0.136088	.0091	9.65	.001	.1829
Children	-0.321327	.0410	5.06	.01	.0277
SES Quintile	0.038911	.0081	3.10	.02	.0142
Anticipations	0.086950	.0216	2.60	.10	.0057
Race	-0.077506	.0299	1.68	---*	.0024
Sex	-0.040627	.0193	1.37	---*	.0019
*Dropped from model.				R-Sq.	.2347
			Adj.	R-Sq.	.2338

NOTES: (1) Standard errors are adjusted in accordance with design effects of the stratified sample used in High School & Beyond. See technical appendix and Skinner, Holt and Smith (1989). (2) Significance level of *t* (p) based on a two-tailed test.

It is not surprising that race and sex fall out of the model, no matter how generous a statistical selection criterion was used[43]. If these variables failed to meet statistical selection criteria at a stage of student history prior to college attendance and in the course of constructing ACRES (see p. 22 above), their chances of playing any role after the student group has been winnowed to 4-year college attendees is dim indeed. Given the nature and intensity of current and continuing concern with the educational attainment of minority students, however, this study keeps race in the models until it fails to register the slightest impact. It is retained in contrast to other features of students' educational histories that are subject to the kind of change (race is not subject to change) that yield degree completion.

If we subjected the specifications for Part B of the Background Model to a logistic regression, would the relationships be the same? Table 30 presents a portrait that, in a few respects is slightly different from that of the OLS accounting. ACRES is again the strongest of the positive components and becoming a parent by age 22 is the strongest of the negative components. The estimates for ACRES say that for each step up the quintile ladder of that variable, the odds of earning a bachelor's degree increase by nearly 100 percent. The estimates for parenthood say that having children reduces the odds of earning a degree by 86 percent (1 minus the odds ratio of 0.14). Degree anticipations and SES are also positive

Table 30—Logistic account of the relationship of pre-college and family variables to bachelor's degree completion among 4-year college students in the High School & Beyond/Sophomore cohort, 1982-1993

Universe: All students for whom the evidence confirms 4-year college attendance at any time (whether transcripts were received or not), who received a high school diploma or equivalent prior to 1988, and who evidenced positive values for all variables in the model. N=4,943. Weighted N=1.179M. Simple s.e.=.697; Taylor series s.e.=1.082; Design effect=1.55.

Variable	Parameter Estimate	Adj. s.e.	t	p	Odds Ratio
INTERCEPT	-3.174	0.691	2.96		
Academic Resources	0.678	0.051	8.58	.001	1.97
Children	-1.968	0.283	4.49	.001	0.14
SES	0.202	0.043	3.03	.01	1.22
Anticipations	0.423	0.112	2.44	.02	1.53
Race	-0.398	0.156	1.65	.10	0.67
Sex	-0.234	0.104	1.08	---	0.79

NOTES: (1) Standard errors are adjusted in accordance with design effects of the stratified sample used in High School & Beyond. See technical appendix and Skinner, Holt and Smith (1989). (2) Significance level of *t* (p) based on a two-tailed test.

contributors (as they were in the OLS version), but in the logistic version both the parameter estimate and the odds ratio favor degree anticipations over SES. The explanation is more technical than substantive. Unlike the OLS version, the logistic account keeps race in the model, however tenuously. This results lends modest support to my decision to carry race into Stage 2.

Model Iteration, Stage 2: Modes of Financial Aid and Student Work

Where do we turn to explore and strengthen the explanation of degree completion? There are four sets of variables to which one's attention is drawn for this task: financial aid and the extent to which a student worked while enrolled, attendance patterns, satisfaction with various aspects of postsecondary education, and first-year performance. Because financial aid has been shown to be critical to initial enrollment decisions (Jackson, 1988) and first-year retention (Stampen and Cabrera, 1986), our story regards it as antecedent to attendance patterns. Financial aid and student work thus enter the model at this point.

The HS&B/So limits our utilization of financial aid variables with confidence. The database relies principally on student self-reports concerning their financing of college education, and students usually don't know where their financial aid is coming from or how much is at issue. For example, the data set includes an unobtrusive Pell Grant file, and I found about 6000 (weighted) students in that file who never claimed—in the surveys—to have received financial aid. Pell Grants, of course, apply to only a fraction of college students; and third-party files for other programs (Stafford Loans, State Supplementary Grants, and others) were not available. A variable was created that simply marked whether a student had ever received a scholarship or grant of any type between 1982 and 1986; and a parallel variable was formulated for loans. With the exception of describing combinations of types of financial aid (including work-study) offered to students (St. John and Noell, 1989), that is the extent of confidence in financial aid analyses for the HS&B/So (and the NLS-72 as well).

The third component of "financial aid" was derived from student responses to a loop of questions concerning the financing of postsecondary education, year-by-year, from 1982 through 1986, with a focus on work. Neither the dollar amount nor the number of hours worked are in question here, rather the fact, sources, timing, and *purpose* of work. For each year students indicated that they worked to pay some of the costs of their postsecondary education a dummy variable was constructed. A positive value was registered if the sources of the student's earnings were Work-Study, Co-Op placements, Teaching/Research Assistantships, and/or "other earnings while in school," i.e. work activities concurrent with enrollment. These activities are more likely to take place on campus than off-campus. Savings from work prior to entering higher education or during periods of stop-out, as well as earnings from summer jobs were excluded. A composite variable was constructed to indicate whether the student "worked" (as defined by these boundaries) in each of the four academic years in question (1982-1986), and then dichotomized to mark whether, in two or more of the four academic years covered by the loop of financing questions, the student worked concurrently with enrollment *and* in order to cover costs of education. The variable is called STUWORK.

The literature on this issue (e.g. Cuccaro-Alamin and Choy, 1998; Horn, 1998) is usually based on the number of hours worked, a datum not available in the HS&B/So. The consensus seems to be that a modest amount of work while enrolled enhances retention, that work on campus certainly intensifies student involvement and contributes to completion (Astin, 1993), but that an excess of work (particularly off-campus) is negatively related to persistence. In the Beginning Postsecondary Students longitudinal studies, NCES posed a rather telling question that conditions the way one must judge student reports of hours worked in relation to persistence and completion. Respondents are asked whether they see themselves primarily as students who happen to be working or employees who happen to be going to school. Table 31, produced from the BPS90 data, confining itself to the first year of enrollment and to a population younger than 24 (to render the data parallel to the HS&B/So universe), provides guidance for interpreting hours worked under that dichotomy. The average hours worked for those who considered themselves students first seems high at 25.9, but it is weighted to the upside by those claiming 40+ hours of work per week and by students attending trade schools.

Table 31.—Hours worked per week during the first year of enrollment (1989-1990), for traditional-aged students in the Beginning Postsecondary Students longitudinal study, by primary role

	Primarily a Student (76% of all)	Primarily an Employee (24% of all)
Proportion employed during >80 percent of months enrolled	52.7% (1.4)	66.0% (3.9)
Proportion of students by range of work hrs./week while enrolled		
1-14	32.4% (1.3)	21.3% (2.0)
15-21	18.0 (1.1)	17.5 (2.2)
22-30	22.8 (1.2)	19.7 (2.1)
31+	26.8 (1.3)	41.5 (2.7)
Total:	100.0%	100.0%
Average work hrs./week while enrolled	25.9 (0.3)	29.9 (0.6)

Note: Standard errors are in parentheses. Source: National Center for Education Statistics: Beginning Postsecondary Students, 1989-94: Data Analysis System.

The contrasts based on primary role would be much greater if the universe had not been confined to traditional-age students (17-24). But even among traditional-age students, one out of four considers himself/herself an employee first, a judgment confirmed by higher average work week hours, proportion employed during most of the months in which they were enrolled, and an obviously higher percentage working full-time (over 30 hours/week).

Because the postsecondary history of the HS&B/So can be as long as 11 years, the role of work in models of bachelor's degree completion by the end of that period may be attenuated. People change status. They become independent; they become employees who happen to be going to school. Combine these changes with growing and swirling patterns of multi-institutional attendance, and Cuccaro-Alamin and Choy's observation that the average work-week of students does *not* change over time, then one doubts that work will emerge as a significant factor in explaining bachelor's degree attainment.

Table 32.—Degree Completion Model: Financial Aid and Employment Variables

Universe: All students for whom the evidence confirms 4-year college attendance at any time, who received a high school diploma or equivalent prior to 1988, and who evidence positive values for the variables in the Background Model. N=4,943; Weighted N=1.179M. Simple s.e.=.697; Taylor series s.e.=1.082; Design effect=1.55.

Variable	Parameter Estimate	Adj. s.e.	t	p	Contribution to R^2
INTERCEPT	1.397895	.1139	7.92		
Academic Resources	0.123885	.0093	8.60	.001	.1829
Children	-0.300230	.0407	4.76	.01	.0277
SES Quintile	0.046395	.0082	3.65	.01	.0142
STUWORK	0.075257	.0209	2.32	.05	.0081
Anticipations	0.079826	.0215	2.40	.05	.0052
Grant-in-Aid	0.067087	.0212	2.04	.10	.0042
Race	-0.083451	.0292	1.84	.10	.0030
Loan	0.019876	.0209	0.61	---*	.0003
				R-Sq.	.2456
				Adj. R-Sq.	.2443

*Dropped from model.

NOTES: (1) Standard errors are adjusted in accordance with design effects of the stratified sample used in High School & Beyond. See technical appendix and Skinner, Holt and Smith (1989). (2) Significance level of *t* (p) based on a two-tailed test.

Indeed, in their initial appearance in table 32, it is obvious that the financial aid/work variables do not strengthen the explanatory power of the model appreciably: the adjusted R^2 increases by barely one percent (.2338 to .2443). Loans (at least the way the data base forces us to define loans) are a very weak addition, while STUWORK is the strongest of the three new variables in terms of its statistical properties in the model. Whether these relationships will hold in the next iteration of the model depends on their interactions with the new variables. Remember that all of the financial aid/work data were derived from student responses to questionnaire items in 1984 and 1986, while the censoring date for bachelor's degree completion is 1993. It is very possible that students received grants and worked while enrolled as undergraduates after 1986, but we don't know. When the calendar for independent variables is truncated and one introduces long-term attendance pattern variables, we might witness changes in the strength of Grant-in-Aid and STUWORK.

Stage 2 of the model (table 32) demonstrates a virtue of carrying forward race, even though the variable did not meet the statistical criteria for retention in the OLS version of the background model. Race is obviously a marginal contributor. Its effects on degree completion for the population of 4-year college students, while slightly negative, are nowhere near the negative effects of having a child by age 22. The more variables in the model, though, the greater the degrees of freedom in statistical analyses, hence the lower the threshold *t* statistic for inclusion.

Model Iteration, Stage 3: Attendance Patterns

The third stage of model development is presented in table 33. This is the first iteration that follows the analysis of attendance patterns, and it changes (and shrinks) the universe. Seven variables were carried forward from the previous iteration. Nine others were introduced here. One of the new variables, number of schools attended, did not meet the $p \leq .2$ selection criterion at all, even in two different formulations: a dichotomous form (one school only v. more than one), and a trichotomous form (one, two, and more than two). All the data we observed in Part III above suggested that this would happen: in an age of multi-institutional attendance, the number of schools attended simply will not be related to degree completion.

The attendance variables brought into play in this iteration can be parsed in three groupings:

(1) Number and order of institutions attended.

NUMSCHL: a simple dichotomy—one school v. more than one.

TRANSFER: a classic pattern in which the student attended a community college first, earned more than 10 credits from the community college, and subsequently earned more than 10 credits from a 4-year college (the "early transfers"–see p. 52 above–are thus excluded); and

NO RETURN: a pattern in which the student attended more than one school, but did *not* return to the first institution of attendance.

(2) Characteristics of the "referent" first institution of attendance.

FIRST4: the first institution of attendance was a 4-year college;

DOCTORAL: the first institution was a doctoral degree-granting institution;

SELECTIVITY: the first institution was selective/highly selective v. non-selective/open door.

(3) Other features of attendance.

NODELAY: the first date of attendance at the first institution occurred 10 or fewer months after high school graduation;

OUTSYS: at some time during his/her undergraduate career, the student attended an institution other than a traditional 4-year or 2-year college; and

NOSTOP: the student was continuously enrolled as an undergraduate.

There are some very dramatic changes in the model with the introduction of attendance pattern variables. The first is that the explanatory power of the model leaps by a factor of nearly 50 percent—from an R^2 of .2456 to one of .3623. This significant advance occurs *not* as a result of the characteristics of the "referent" institution of attendance, rather as a by-product of student movement among institutions and the temporal dimensions of enrollment.

The most notable changes in this iteration of the model are the dominance of continuous enrollment (NOSTOP), the contracting strength of the pre-collegiate variables (ACRES and SES Quintile), and the superficially contradictory positions of No Return and Transfer. This apparent contradiction is fairly easy to explain. In the No-Return variable, the student attends more than one institution and does NOT return to the first, and this behavior has a negative relationship to degree completion. As defined, Transfer involves a No Return-*type* situation, but with a specific sequence and criteria (the student must earn more than 10 credits in a 2-year college *before* earning more than 10 credits in a 4-year college). Transfer has a positive relationship to degree completion. Students in a classic transfer pattern are moving *toward* a bachelor's degree. Students in No-Return positions include *reverse* transfers who move *away* from the path toward the bachelor's degree.

On the other hand, *selectivity* of the referent first institution of attendance has a significant role to play, even though its statistical position is tenuous. Covariance analysis provides some clues as to the lower standing of selectivity in the model: it is entangled with the pre-collegiate variables ACRES and Anticipations, the former being strong enough to reduce the

Table 33.—Degree Completion Model; Attendance Pattern Iteration

Universe: All students who attended a 4-year college at any time, whose transcript records were complete, and who evidenced positive values for all variables in the model. N=4,538. Weighted N=1.079M. Simple s.e.=.714; Taylor series s.e.=1.087; Design effect=1.52.

Variable	Parameter Estimate	Adj. s.e.	t	p	Contribution to R^2
INTERCEPT	0.610047	.1910	2.10		
NOSTOP	0.323020	.0204	10.42	.001	.1994
Academic Resources	0.091391	.0090	6.68	.001	.0989
Children	-0.215601	.0413	3.43	.01	.0135
Transfer	0.231072	.0313	4.86	.001	.0115
No Return	-0.106825	.0215	3.27	.01	.0106
SES Quintile	0.030959	.0078	2.61	.02	.0091
Grant-in-Aid	0.064003	.0185	2.28	.05	.0053
Anticipations	0.050385	.0206	1.61	---*	.0036
Selectivity	0.082775	.0259	2.10	.05	.0031
Race	-0.081385	.0276	1.94	.10	.0028
STUWORK	0.051003	.0191	1.76	.10	.0023
FIRST4	0.063635	.0319	1.31	---*	.0023
OUTSYS	-0.080847	.0530	1.00	---*	.0008
No Delay	0.047080	.0294	1.05	---*	.0008
DOCTORAL	0.025436	.0204	0.82	---*	.0005
*Dropped from model.				R-Sq.	.3644
			Adj.	R-Sq.	.3623

NOTES: (1) Standard errors are adjusted in accordance with design effects of the stratified sample used in High School & Beyond. See technical appendix. (2) Significance level of *t* (p) based on a two-tailed test.

effects of selectivity. This is another case of common sense: one assumes that most students who start out in selective or highly selective institutions take an academically intense curriculum in high school, perform decently (if not very well) on SATs and ACTs, stand toward the top of their classes., and are committed to earning at least bachelor's degrees. Among other place-related variables, starting in a doctoral degree-granting institution and stepping outside the secular higher education system were so weak as to fall below minimum significance.

Probably more shocking to conventional analysis, however, are the elimination of No-Delay and Anticipations from the model. Remember, though, just how much the definition of these variables departs from those of conventional analysis. No-Delay allows into the model students who did not graduate from high school on time but who enrolled in postsecondary education within 10 months of their graduation date, and not as a false start. These students will attenuate the effects of a traditional "delayed entry" account. While the Anticipation variable appears to contribute more to the explanatory power of the model than other marginal independent variables such as selectivity of first institution, race, and STUWORK, it did not meet the *t* statistic threshold criterion.

Model Iteration, Stage 4: Performance in the First "True" Year

The next group of variables to enter in this model involves various configurations of academic performance in the *referent* first year of attendance. The justification for focusing on this group of variables lies in a tradition of the literature emphasizing the critical role of freshman year performance in retention. The literature, though, unfortunately relies on student self-reports of grades, and without reference to credits attempted or earned, and is not very rigorous concerning what it means by the "freshman year" (e.g. Kanoy, Wester and Latta, 1989). There are exceptions (e.g. Smith, 1992), but they are rare.

The "referent first year" performance variables used in this study are:

- Freshman GPA. GPAs were determined for the first full-year of attendance, and were set out in quintiles. Freshman GPA is a dummy variable that divides performance in the top two quintiles from the bottom three. For the HS&B/ Sophomore cohort, the dividing line turns out to be 2.70.

- Credit Ratio. The ratio of credits earned to credits attempted during the first "true" year of attendance. Students who earned less than 90 percent of attempted credits were declared on one side of a dichotomy. But students who attempted 10 or fewer credits were excluded from this calculation.

- Low-Credits. Students who earned less than 20 credits in their first "true" year of attendance stand on one side of a benchmark. Some of these students did not *attempt* more than 20 credits. A combination of Credit-Ratio and Low-Credits was tested as a proxy for part-time status in the first true year. It was not as convincing as the DWI Index described above (see p. 55-56), but DWI is derived from an entire undergraduate career and hence is not part of the "true first year" performance variable configuration.

Table 34 brings these variables into the model. Compared to the population in the previous iteration, we lose a weighted N of about 50,000 students. We lose some of them because as soon as the independent variables are based on grades, credits, and a distinct time period, we can no longer include students with missing transcripts, even if we know a great deal about the institutions of those missing documents. We also lose those students whose entire first "true" year of attendance was consumed with non-credit, no grade remedial courses. Calculations for any of the three first year attendance variables for these obviously weaker students are impossible under those conditions.

The loss of these students inevitably skews the model. Weaker students don't finish degrees. The large group that remains in the equation will thus exhibit less variance in degree completion. The new contracted universe thus slows down the increase in the explanatory power of the model: the adjusted R^2 moves by less than one percent (from .3623 in the stage 3 iteration to .3696).

The freshman year performance variables have a significant impact on both the composition of the model and the relative weights of the remaining attendance pattern variables. Selectivity of first institution of attendance, financial aid in the form of scholarships or grants-in-aid, and, finally, race, fall out of the model altogether. Race falls out of the model for the same reason that the contribution of SES declines even more from the Attendance Pattern iteration (stage 3): as one moves across the college access line, across the 4-year college attendance line, and into course-taking and academic performance, demographic variables are less and less important. What you do becomes more and more important than where you came from, though the effects of SES will never wholly be washed away. The last of the characteristics of first institution of attendance to remain in the model, selectivity, also falls away for an analogous reason: what you do is more important than where you are.

But the case of Grant-in-Aid calls for explanation. In the Attendance Pattern iteration (table 33), its contribution to the explanatory factor (the R^2) was larger than that of STUWORK. Now the positions are reversed, and the student work variable survives, while the scholarship variable does not. Why? The answer is a by-product of the way STUWORK was defined as a dichotomous variable: the positive value was assigned only when the student worked while enrolled in college for *more than one* of the first four years following scheduled high school graduation in 1982, whereas a positive value was assigned to Grant-in-Aid if the student received a grant-in-aid or scholarship in *any one year* during the same period. Persistence is thus implicit in STUWORK, not in Grant-in-Aid.

The most significant change between the Attendance Pattern and First Year Performance models, though, is that the "academic resources" (ACRES) variable comes back to the top of the list, changing places with continuous enrollment (NOSTOP). The fact that the academic resources brought forward from secondary school comes to play such a robust role in the model only confirms the power of true first year performance: strong academic resources provide students with momentum into their college years--whenever those years start. Of the other "family" background variables, the contribution of "children" has been unaffected in any of the iterations of this model, indicating that if one starts a family at a young age, one's chances of completing a bachelor's degree by age 30 are negatively affected no matter what one does (see Waite and Moore, 1978).

Table 34.—Degree Completion Model: First Year Performance Indicators

Universe: All students who attended a 4-year college at any time, whose transcript records were complete, and who evidenced positive values for all variables in the model. N=4,264. Weighted N=1.029M. Simple s.e.=.723; Taylor series s.e.=1.079; Design effect=1.49.

Variable	Parameter Estimate	Adj. s.e.	t	p	Contribution to R^2
INTERCEPT	1.391811	.1547	6.04		
Academic Resources	0.084774	.0087	6.54	.001	.1668
NOSTOP	0.237503	.0213	7.48	.001	.1017
Low Credits	-0.183009	.0268	4.58	.001	.0351
Freshman GPA	0.096158	.0190	3.40	.01	.0154
Children	-0.250517	.0425	3.96	.01	.0152
Transfer	0.158913	.0238	4.48	.001	.0099
No-Return	-0.109283	.0211	3.48	.01	.0096
SES Quintile	0.033787	.0075	3.02	.02	.0076
Credit Ratio	-0.097777	.0316	2.08	.10	.0036
STUWORK	0.047759	.0184	1.74	.10	.0026
Selective	0.062054	.0251	1.66	---*	.0019
Grant-in-Aid	0.035557	.0184	1.30	---*	.0010
Race	-0.048902	.0277	1.18	---*	.0009

*Dropped from model.

R-Sq. .3715
Adj. R-Sq. .3696

NOTES: (1) Standard errors are adjusted in accordance with design effects of the stratified sample used in High School & Beyond. See technical appendix. (2) Significance level of *t* (p) determined by a two-tailed test.

As suspected, the ratio of credits earned to credits attempted in the first true year of attendance might have been a stronger contributor to the model had I found a way to include remedial students who attempted no additive credits during that period. But the presence of Low Credits, reflecting both part-time students and those who earned less than 90 percent of the credits attempted, is a far better explanation for the marginal impact of the credit ratio. Low-Credits is an umbrella for behaviors that hamper momentum toward degrees. As such, it will suppress variables expressing sub-sets of those behaviors.

If any of the variables describing referent institution of first attendance had been retained in the model through this iteration of true first year performance, we would be invited to dig further into institutional characteristics looking for significant effects. Roughly 30 percent of the

HS&B/So postsecondary students who earned more than 10 credits and attended a 4-year college at some time also attended more than one school *and* never returned to the first institution of attendance. For this group of students, institutional effects are very difficult to attribute, even if their institutions were campuses in the same system[44]. Another 12 percent attended three or more institutions but returned to the first. These groups constitute such a large proportion of 4-year college students as to give pause to any institutional effects analyses (e.g. Velez, 1985) in national data bases.

Model Iteration, Stage 5: Continuing Effects of the First Year

But other variables that can be derived from the rich archives of NCES longitudinal studies suggest that student experience at the first institution of attendance may have continuing effects. The extensions of three performance variables encourage us to explore this hypothesis. After all, as one moves beyond the first true year of attendance, the model can—and should—be used to guide post-matriculation advisement. To fill the tool box with appropriate instruments and directions, we should transcend the boundaries of traditional prediction models.

DWI Again

One of these variables is the "DWI Index" (see pp. 55-56). When students withdraw or leave incomplete a significant percentage of their attempted courses, the behavior is bound to have a negative effect on degree completion. In table 37, DWI is a dummy variable with a cut-point of 20 percent. That is, the variable is positive if students dropped, withdrew from, left incomplete, or repeated more than 1 out of 5 courses attempted during their undergraduate careers. The correlation of DWI with the first-year variable, Low-Credits, is unsurprisingly high in a sample of this size and in a matrix with eleven other variables (.233; $t=10.3$). If you are withdrawing from a significant number of courses, the chances are reasonably high that your credit count will be low, no matter which year of your undergraduate career is at issue.

GPA Trend

The second variable that extends first year history is the *trend* of a student's GPA. First true year grade performance, as we have seen, plays a modestly positive role in the model of degree completion. But common sense tells us that for some students, first year grades will be *lower* than final GPA for people who complete degrees. Students begin to major, and grades in major courses are inevitably higher than grades in prerequisites or distribution requirements. The variable GPA Trend is a dichotomous version of the ratio of final year GPA to first year GPA. If the range of that ratio was .95 to 1.05, there was basically no change in performance. Below .95 indicates a falling GPA; above 1.05 reflects a rising GPA. GPA-Trend places a positive value on the rising GPA, and a negative value on any ratio of .95 or lower.

How did the HS&B/So students who attended 4-year colleges at any time fare in each of these trends in terms of final GPA? Table 35 tells an interesting story in this matter. Among those who earn bachelor's degrees, students whose GPAs don't change that much between first and final undergraduate years sport a higher final GPA than those whose GPAs rise. But the bachelor's degree attainment rate of those with "stable" GPAs is lower than that of those whose grades rise over time. Doing better (with a rising GPA as the indicator) appears to be an indirect proxy for determination.

Table 35.—Final undergraduate GPA of students who attended 4-year colleges at any time, by trend in undergraduate GPA and bachelor's degree attainment, High School & Beyond/Sophomore cohort, 1982-1993

	Final GPA of Students Who Did Not Earn Bachelor's			Final GPA of Students Who Earned Bachelor's			% of Trend Group Who Earned BA	**Trend Pct. of All**
	Mean	S.D.	s.e.	Mean	S.D.	s.e.		
GPA TREND:								
Rising	2.39	0.651	.0246	2.87	0.442	.0110	69.8%	**41.1%**
Stable	2.32	0.800	.0031	3.02	0.485	.0014	64.1	**34.1**
Falling	1.97	0.599	.0023	2.70	0.429	.0016	52.2	**24.9**

NOTES: (1) Universe consists of all HS&B/So students who attended a 4-year college at any time and for whom an undergraduate GPA could be computed. Weighted N=1.25M. (2) Standard errors adjusted for design effect of 1.49. (3) Differences in bachelor's attainment rates are significant at $p \leq .05$. **SOURCE**: National Center for Education Statistics: High School & Beyond/Sophomore cohort, NCES CD#98-135.

Remedial Problems

A third variable seeks to add the effects of remedial problems. The college transcript samples of NCES longitudinal studies teach us that there are different kinds of remedial course work, and that some are more serious than others. If the *type* of remediation matters, so does the amount. Of the HS&B/So students who were assigned to remedial *reading*, 74 percent were enrolled in two other remedial courses. Of those whose only remedial mathematics work in college was pre-collegiate algebra, only 16 percent were enrolled in two or more remedial courses (and of this group, 75 percent were assigned to remedial reading). The first case is a remedial problem case; the second is not. These are further matters of common sense: (1) people with reading deficiencies cannot read mathematics problems, either; (2) people whose only problem on a basic skills placement test stems from a bad Algebra 2 course in high school can proceed toward a degree with minimal disruption.

The "remedial problem" variable used in the regression model is a trichotomy derived from the observations of table 36 on the relationship between remedial coursework and degree completion among those who attended a 4-year college at any time: 1=any remedial reading, 2=other types of remedial work, and 3=no remedial work.

Table 36.—Bachelor's degree attainment of 4-year college students with different types and amounts of remedial coursework, High School & Beyond/Sophomore cohort, 1982-1993

	Percent of All Students	Percent Earning Bachelor's Degree
Any remedial reading	10.2	39.3
No remedial reading, but >2 other remedial courses	18.7	46.5
No remedial reading, but 1 or 2 other remedial courses	20.4	59.6
No remedial coursework	50.7	68.9

NOTES: (1) Universe consists of students who attended a 4-year college at any time and for whom transcript data on remedial coursework were available; Weighted N=1.38M; (2) all column pair comparisons are significant at $p \leq .05$.; (3) For the definition of remedial courses, see footnote #8. **SOURCE**: National Center for Education Statistics: High School & Beyond/Sophomore cohort, NCES CD#98-135.

The "remedial problem" variable is treated as a continuing effect, and not a first year performance variable, for a very empirical reason. Among the HS&B/So students who attended a 4-year college at any time *and* took one or more remedial courses, slightly more than half of those courses (52.2 percent) were taken during the first calendar year of attendance. In fact, by the end of the second calendar year of attendance 68.6 percent of the total 11-year remedial course load had been completed; and by the end of the fourth year, 84.4 percent. These data suggest that to isolate the impact of remediation problems on degree completion, we should look beyond the first year.

Table 37 presents the "extended first year performance" iteration of the regression model in which these three variables are introduced. It is not surprising that first year Credit Ratio is pushed out of the model by DWI and GPA Trend. The DWI index will also reduce the contribution of Low Credits (from .0382 to .0262) because it overlaps Low-Credits (the correlation of .2396 in a matrix with a dozen other variables is strong). The remaining

Table 37.—Degree Completion Model: Extending First Year Performance

Universe: All students who attended a 4-year college at any time, whose transcript records were complete, and who evidenced positive values for all variables in the model. N=4,264. Weighted N=1.029M. Simple s.e.=.723; Taylor series s.e.=1.079; Design effect=1.49.

Variable	Parameter Estimate	Adj. s.e.	t	p	Contribution to R^2
INTERCEPT	1.616038	.1409	7.70		
Academic Resources	0.077990	.0081	6.46	.001	.1668
NOSTOP	0.206917	.0204	6.81	.001	.1017
DWI Index	-0.254825	.0249	6.87	.001	.0642
Low-Credits	-0.178315	.0256	6.97	.001	.0225
Freshman GPA	0.141591	.0190	5.00	.001	.0225
GPA Trend	0.155576	.0179	5.83	.001	.0153
Children	-0.229453	.0405	3.80	.01	.0119
No-Return	-0.096415	.0201	3.22	.01	.0071
Transfer	0.129201	.0227	3.82	.01	.0068
SES Quintile	0.029912	.0069	2.91	.02	.0053
STUWORK	0.049976	.0176	1.91	.10	.0026
Credit Ratio	-0.060221	.0300	1.35	---*	.0018
Remedial Problem	-0.019793	.0139	0.96	---*	.0006
*Dropped from model				R-Sq.	.4291
			Adj.	R-Sq.	.4275

NOTES: (1) Standard errors are adjusted in accordance with design effects of the stratified sample used in High School & Beyond. See technical appendix. (2) Significance level of *t* (p) determined by a two-tailed test. (3) DWI=Drops, Withdrawals, and Incompletes.

attendance pattern variables, No-Return and Transfer, also decline in influence in the face of the new variables built on grades. In other words, after the first year of attendance, academic performance (with grades as its proxy) becomes more important for degree completion than place-referenced attendance.

The new Remedial Problem variable fails all tests for entry into the model. Why? In a way, whatever remedial variable might be introduced at this point is doomed *a priori* by the strength of the secondary school Academic Resources (ACRES) variable. The correlation between ACRES and Remedial Problem, in the matrix with a dozen other variables, is a stunningly high -.4295. Common sense wins again: students entering college with a low degree of academic resources evidence continuing remedial problems dominated by reading and do not earn degrees. ACRES already accounts for this story, as do other variables such as DWI and Low-Credits, which are inclusive in their reflection of sub-par academic work.

More significantly is that the influence of SES begins to fall (from .007 in table 34 to .0053 in table 37), and that the contributions of continuous enrollment (NOSTOP) and academic resources (ACRES), are unaffected by the introduction of post-first year performance variables. These two variables keep the fundamental story alive.

Model Iteration, Stage 6: Can We Get Satisfaction?

If only three attendance pattern variables remain in the model (continuous enrollment, No-Return, and transfer), and the model accounts for about 43 percent of the variance in degree completion, what's left to add? How much further can we push the envelope of explanation of bachelor's degree completion for the HS&B/So cohort? An R^2 of .4275 in a model such as this is considered extraordinarily persuasive. Some would argue to stop at this point, that the models have reached a plateau of explanation. Others would note that in the more traditional terms of predictive modeling, we reached a plateau of explanation in Stage 4 (first year performance). But there is a substantial body of research concerned with student responses to, and assessments of their postsecondary experience that suggests these responses might influence persistence (e.g. Tracey and Sedlacek, 1987; Pascarella, Terenzini, and Wolfle, 1986). Too, in trying to replicate Horn's (1998) analysis of the BPS90 students who left postsecondary education during or by the end of their first year, and exploring what distinguished those who returned at a later point in time (the stop-outs) from those who never returned (what Horn calls the "stay-outs"), I found an almost linear relationship between degree of dissatisfaction and permanent "stay-out" status. The group at issue was too small to yield significant findings, but the experience suggested that the construct of satisfaction might be profitably pursued in a multivariate context.

The fifth collection of variables used in the iteration covers four aspects of student satisfaction with their postsecondary careers: academic, environmental, work preparation, and cost. The questions were asked only once, retrospectively, in 1986 (four years after scheduled high school graduation). Of the four categories, only two--academic and environmental--offer enough items to create a separate index[45]. For each item, e.g. satisfaction with "my intellectual growth" (an academic category) or "sports and recreation facilities" (an environmental category), students were offered a scale of five responses. I turned each question into a dummy variable (dissatisfied/not dissatisfied), aggregated the responses, and turned each of the aggregates into another dummy variable. A composite "dissatisfaction index" was then built from all four categories with a minimum score of 4 (highly satisfied) to a maximum of 8 (highly dissatisfied). The composite was again dichotomized, with scores from 6 to 8 signifying some degree of overall dissatisfaction.

Only one question about satisfaction with the costs of postsecondary education was asked. To isolate the contribution of financial aid to this dimension of satisfaction, an enhanced dummy variable was created[46]. Students were assigned a positive value if they received either a grant or a loan *and* were dissatisfied with the costs of postsecondary education. The components of this variable were first treated separately, and the difference in dissatisfaction rates between those with grants and those with loans was found to be small (30.8 percent to 33.6 percent) and statistically insignificant.

One reviewer of this study suggested that one should turn the coin, so to speak, on these features, and emphasize satisfaction, not dissatisfaction. Given dictotomous variables, too, there is a statistical argument for choosing the larger group (those who are satisfied) as the reference point. But there is substantial body of research—Horn's (1998) included—demonstrating that indications of global satisfaction are almost mindless reflex responses of students, and that, if one wishes to isolate a strong attitude in an explanatory context, the negative attitude will be more revealing. People responding to surveys have to go out of their way to tell you that they are unhappy.

We are unfortunately limited by the data base in our grasp of those aspects of student experience that Astin (1977, 1989, 1993) and Tinto (1975, 1993) and others have described in such terms as "academic integration" and "social integration," though, as Cabrera, Nora and Castaneda (1993) demonstrate, academic integration is expressed indirectly through GPA. High School & Beyond never asked (as does the Beginning Postsecondary Students, 1989-94 study) how much contact with faculty students enjoyed outside of class, for example. And with the exception of athletics, we have a very limited sense of their participation in extra-curricular activities in college. These questions may be important, but when students are attending two or three schools, and when the fact of multi-institutional attendance doesn't seem to matter in explaining degree completion, then it is impossible to attach the academic and social experiences elicited by these questions to any one institution, unless the student's career involved only one institution (see Cabrera, Casteneda, Nora, and Hengstler, 1992), or, if more than one, was dominated by an institution at which one began and to which one returned.

Table 38 presents the results of the stage 6 iteration of the model. We lose a few people here as a by-product of non-response to the satisfaction questions, and the design effect drops from 1.49 to 1.46, resulting in slightly higher critical *t*-values than would have been the case with a larger sample. The input variables include the eleven carried forward from stage 5 (Extending First-Year Performance), three satisfaction indicators, and the combination financial/satisfaction dummy variable. Given the degree of overlap in the satisfaction variables (confirmed by covariance analysis), the selection set a generous inclusion threshold of $p \leq .2$ if for no other reason than to demonstrate just how marginal some of these variables would be. Even then, only one of the four satisfaction variables, that indicating dissatisfaction with *academic experiences*, passed the test—only to be dropped within the dynamics of the regression.

Table 38.—Degree Completion Model: Satisfaction Variables

Universe: All students who attended a 4-year college at any time, whose transcript records were complete, and who evidenced positive values for all variables in the model. N=3,807. Weighted N=914k. Simple s.e.=.741; Taylor series s.e.=1.083; Design effect=1.46.

Variable	Parameter Estimate	Adj. s.e.	t	p	Contribution to R^2
INTERCEPT	1.688707	.1448	7.99		
NOSTOP	0.210741	.0213	6.78	.001	.1581
Academic Resources	0.079632	.0086	6.34	.001	.1011
DWI* Index	-0.272824	.0259	7.22	.001	.0714
Low-Credits	-0.192789	.0252	5.24	.01	.0207
GPA-Trend	0.125959	.0182	4.74	.01	.0173
Children	-0.230705	.0431	3.67	.01	.0136
Freshman GPA	0.133546	.0193	4.74	.01	.0087
No-Return	-0.105192	.0205	3,52	.01	.0087
Transfer	0.120377	.0228	3.62	.01	.0057
SES Quintile	0.027110	.0070	2.65	.05	.0045
STUWORK	0.042087	.0176	1.64	---*	.0018
Acad Dissatisfaction	-0.035955	.0198	1.24	---*	.0011
*Dropped from model.				R-Sq.	.4128
			Adj.	R-Sq.	.4109

NOTES: (1) Standard errors are adjusted in accordance with design effects of the stratified sample used in High School & Beyond. See technical appendix. (2) Significance level of *t* (p) determined by a two-tailed test. (3) *DWI=Drops, Withdrawals and Incompletes.

It is obvious that, after the first true year of higher education, the financial aid and satisfaction factors are peripheral. The overall explanatory power of the model actually decreases with the interactions of the satisfaction measures. They have the effect of reversing the positions of ACRES and NOSTOP again, and slightly reducing the contribution of both those key variables. The satisfaction index proved to be unproductive, as did the attempt to tease out dissatisfaction with the costs of higher education. Very little of this was a result of reducing the universe of students by eliminating those who did not answer any or enough satisfaction questions in 1986 to yield a satisfaction index score. Had we not eliminated those students, there would have been virtually no change in the R^2 indicator.

Reintroducing Race, Gender and Aspirations

If there were aspects of the fundamental story reflected in table 37 that differed significantly by race and sex, the reintroduction of those variables at this point in model-construction would alter the relationships we have observed. Following Astin, Tsui and Avalos (1996), I made one more attempt to bring them into the model, starting with the variables that survived Stage 5 (Extending First-Year Performance) and this time with four separate race variables (white v. others; black v. others; Latino v. others; and Asian v. others). None of the race variables was accepted into the equation, and the effect of the attempt basically left the R^2 unchanged at .4263.

But the modified aspirations variable (Anticipations) did pass the threshold criterion of statistical significance, reentering the model, and moving ahead of SES Quintile and STUWORK in its contribution to the explanatory power of the model. While carrying a comparatively low *t* (2.29), and a low degree of significance (.10), its reintroduction slightly boosts the adjusted R^2 from .4275 to .4309. In light of the fact that all regression analysis involves some degree of multicollinearity (Schroeder, Sjoquist, and Stephan, 1986), and the Anticipations variable evidences a high degree of correlation with the performance variables in the model, I wouldn't make that much of its reappearance. It can move in and out of a model such as this, but always on the margin. Its minor contribution at this point indicates that, in fact, a plateau of explanation was reached in the first-year performance extension iteration. Given the dependent variable, bachelor's degree *completion*, anything else would muddle and distort the guidance this story offers.

Confirming the Guidance: the Logistic Story

The preferred statistical technique for telling this story involves logistic regression. To put the difference between the Ordinary Least Squares linear regression models and the logistic model too simply, the former seeks to minimize the errors in the measurement of an event, while the latter seeks to estimate the maximum likelihood of an event.

Table 39 takes the first five equations from the OLS story–that is, up to the point at which the story reached a plateau–and presents them in a logistic form. By displaying the equations together, we can both observe the changes in the position of the independent variables from step to step, and assess whether the sequence of models provides an increasingly convincing explanation ("goodness-of-fit"). What do we see in Table 39? First, there are five statistics at the bottom of page 81 that help us judge the explanatory power of the sequence of models and that tell us which stages make the greatest difference. To put it simply, everything that is supposed to happen in a series of logistic models such as this (Cabrera, 1994), happens: the G^2 declines; the ratio of G^2 to the degrees of freedom declines; the Chi-square rises in the face of an increasing number of variables. All of this says that the logistic analysis becomes increasingly efficient. The relative changes of these statistics also tell us that the Financial Aid model does not add that much to the potential guidances in our tool box, and that the Attendance Pattern model adds the most (the same conclusions reached in the OLS sequence).

To appreciate the convergences and differences between logistic and linear models of this story, one must watch the changes in the odds ratios for a given variable and the statistical significance of its parameter estimates. For example, the variable indicating that the first college attended was a 4-year institution exhibits seemingly impressive odds ratios in all three of its appearances, but only in one of those cases is the estimate statistically significant, and

Table 39.-Logistic version of 5 sequential regression models with bachelor's degree attainment by age 30 as the outcome. High School & Beyond/Sophomore Cohort, 1982-1993

	1: Background		2:Financial Aid		3: Attendance Patts		4:1st Year Perform		5.Extending Perform	
	Estimate	Odds Ratio	Estimate	Odds Ratio	Estimate	Odds Ratio	Estimate	Odds Ratio	Estimate	Odds Ratio
Intercept	-3.174		-3.248		-4.212		-3.609		-3.548	
ACRES	0.678*	1.97	0.628*	1.87	0.544*	1.72	0.403*	1.50	0.389*	1.48
Children	-1.968*	0.14	-1.924*	0.15	-1.469**	0.23	-1.517**	0.22	-1.480**	0.23
SES	0.202**	1.22	0.253**	1.29	0.194***	1.21	0.162++	1.18	0.161++	1.18
Anticipations	0.423***	1.53	0.386+	1.47	0.226	1.25	0.322	1.38	0.478+	1.61
Race	-0.398++	0.67	0.461++	0.63	0.461++	0.63	-0.325	0.72	-0.069	0.93
Sex	-0.234	0.79	-0.201	0.82	-0.150	0.86	-0.091	0.91	0.006	1.01
Grant-in-Aid			0.370+	1.44	0.405+	1.50	0.215	1.24	0.201	1.22
Loan			0.134	1.11	0.056	1.06	0.059	1.06	0.044	1.05
STUWORK			0.396++	1.49	0.259	1.30	0.200	1.22	0.228	1.26
# of Schools					-0.009	0.99	0.015	1.02	-0.077	0.93
First Was 4-Yr					0.403	1.50	0.603++	1.83	0.441	1.55
First Was Doctoral					0.153	1.17	0.187	1.21	0.171	1.19
First Was Select					0.693+	2.00	0.600++	1.82	0.561	1.75
OUTSYS					-0.652	0.52	-0.564	0.56	-0.782	0.46

Transfer					1.349*	3.85	1.449*	4.26	1.255*	3.51
No Return					-0.691**	0.50	-0.634**	0.53	-0.522+	0.59
No Delay					0.262	1.30	0.166	1.18	0.214	1.24
No Stop					1.553*	4.73	1.408*	4.09	1.349*	3.85
1st Yr Grades							0.726*	2.07	1.104*	3.02
1st Yr Low Creds							-0.887*	0.41	-0.927**	0.40
1st Yr Cred Ratio							-0.573++	0.56	-0.501	0.61
DWI Index#									-1.558*	0.21
Grade Trend									1.167*	3.21
Remedial Problem									-0.137	0.87
G^2	5386.5		5304.6		4106.7		3643.1		3300.3	
df	4929		4926		4419		4237		4234	
G^2/df	1.0928		1.0768		0.9293		0.8598		0.7795	
X^2 (df)	35.61 (6)		36.74 (9)		41.15 (18)		41.64 (21)		45.57 (24)	
p	.001		.001		.01		.01		.01	

NOTES: (1) The universe for each stage in the model is the same as that used for the parallel steps in the Ordinary Least Squares regressions above. (2) Standard errors used in the determination of statistical significant of the Beta estimates are adjusted by the same design effects as in the OLS versions. (3) Keys to significance levels: *=.001 **=.01 ***=.02 +=.05 ++=.10. (4) #DWI Index=Drops, Withdrawals and Incompletes. SOURCE: National Center for Education Statistics: High School & Beyond/Sophomore cohort. NCES CD#98-135.

even in that case, barely. The linear version did not even allow the "First Was 4-Year" variable into its models. On the other hand, in the logistic version, SES evidences more modest odds ratios, but its estimates are statistically significant in all five appearances.

The logistic story, unlike the linear story, truly disentangles the Transfer and No Return variables. There is a dramatic diversion between the two, and Transfer, in particular, turns out to be much stronger in the "maximum likelihood" approach to bachelor's degree completion than it does in the linear model. The Transfer variable is very distinct in this account. It does *not* mean merely that you attended both a 2-year and a 4-year college, rather that you started at the 2-year, earned more than 10 credits, then moved to the 4-year and earned more than 10 credits there, too. This definition truly sorts people moving toward a bachelor's degree from those multi-institutional attendees engaged in less direct routes. It also filters out the "early transfers," those who did not wait for the community college to provide a sufficient comfort level in higher education. As noted above (p. 52) those who jump ship early to the 4-year college are much less likely to complete a bachelor's degree. The odds ratios for Transfer so defined are very high: 3.85:1 in the Attendance Pattern model, 4.26:1 when we add the 1st Year Performance variables, and 3.51:1 in the Extended Performance model. The only other variable in the logistic story that exceeds those odds ratios is NOSTOP, that is, continuous enrollment. The transfer focus thus becomes a critical direction in the tool box. If we know that students who meet the transfer sequence criteria succeed as well as they do, then we should guide them into that sequence instead of allowing them to leave the community college too early.

What else is different when we compare the logistic to the linear account? The Academic Resources variable, while statistically significant in all its appearances in the logistic model, exhibits an odds ratio decline to the point at which its power appears to be less than first year grades, overall GPA trend, Transfer, and continuous enrollment. The power is still considerable, but not as overwhelming as the linear story would lead us to believe. While this is a disappointment to the tenor of the analysis to date, the most significant drop in the odds ratio for ACRES occurs in the 1st Year Performance iteration (Stage 4), when the restricted Transfer variable enters. Otherwise, the two stories–linear and logistic–are the same, and the same 11 variables remain statistically significant in the final step of the model.

V. Conclusion: the Tool Box Story

The story told by this voyage is clear, and advances the observation of Alexander, Riordan, Fennessey, and Pallas (1982) that "academic variables are much more potent predictors of college completion" than social background variables (p. 324), Hearn's analysis of "non-traditional enrollment styles," and Cabrera, Nora, and Castaneda's (1993) remarkable adjustment in the retention-modeling literature in unburying academic performance as a powerful, direct influence on the momentum toward degrees. It helps us advise and guide students no matter what paths of attendance they follow through higher education. It tells us that if degree-completion lags for any student or group of students, the situation is fixable. We learn where to take the tool box, and what tools to use.

One must acknowledge that SES has a continuing influence in life-course events. But the analysis here (and elsewhere) indicates how much education can mitigate those effects–and in both directions (downward mobility is not a chimera). If SES were an overpowering presence, the tool box directions would be futile. Optimism is the preferred stance.

There is an obvious academic line to the story. To admit that some students are not prepared for the academic demands of the particular higher education environment in which they initially find themselves seems to be some form of heresy in the research traditions on this issue. We do students no favors by advancing to center stage complex stories built from marginal variables to explain why they don't finish degrees. Yes, there are psychological and personal reasons for non-completion, but these are extraordinarily difficult to micromanage. The tool box does not offer much help. At the same time, "academic demands" and "environment" are both relative and varied. Some students may struggle to major in a field they did not understand well enough before they cross the threshold of the major, e.g. engineering (Adelman, 1998) or over-estimate their own talent and proclivities in the field, e.g. music. If their academic performance lags, they may (1) migrate to another major either without stopping-out or with a short term of non-continuous enrollment, (2) stop-out and rethink what they are doing in higher education, perhaps returning in one or two years, (3) turn up as provisional permanent drop-outs at age 30, or (4) change institutions in combination with changed field and stop-out. As this analysis has demonstrated, changing institutions has minor effects on completion. Academic preparation, continuous enrollment, and early academic performance, on the other hand, prove to be what counts.

This tripartite message has dimensions that can—and should—be heeded by those to make and execute policy for higher education, by those who advise and guide students in both high school and college, by those involved in research and evaluation of the school/college continuum, and by students themselves. Let's reflect on these dimensions.

Start with Opportunity-to-Learn

Think about the ACRES variable, the foundation of this analysis. It has three components, only *one* of which is subject to change, the curriculum component. For purposes of constructing the ACRES index, curriculum intensity and quality was set out as a scaled continuum. But every point on that scale can be described as a criterion, a *standard of content*. While test scores and class rank reflect standards of *performance* that are usually expressed in relative terms, there is no reason why all students cannot reach the highest levels of the curriculum scale—at which point, of course, the scale itself can be abandoned. This ideal state will not come to pass without ensuring opportunity-to-learn, and, at the present moment, not all secondary schools can—or do, or will—provide that opportunity (Spade, Columba, and Vanfossen, 1997). Many do not offer mathematics beyond Algebra 2; many offer Algebra 2 courses that, in content, are closer to Algebra 1. Many cannot offer the three basic laboratory sciences, or foreign language study beyond the second year, or computer programming–let alone Advanced Placement courses. Students who enter higher education from these schools, enter with less momentum toward degrees than others. Poor and working-class students, students from rural areas, and minority students are disproportionately affected by this lack of opportunity-to-learn (Rosenbaum, 1976; Monk and Haller, 1993).

What can the higher education enterprise do to provide equitable opportunity-to-learn? Dual enrollment, a growing practice, is one answer. The higher education partner in this arrangement is usually a community college. Under dual enrollment, high school students who do not have access to trigonometry or physics or third year Spanish at the high school take those courses at the local community college and receive both high school and college credit for them. Direct provision can also fill curricular gaps at those high school districts willing to accept college faculty who provide the instruction on site. It's nigh impossible to do this in laboratory science if there is no laboratory in the secondary school, but a variety of other subjects are open to regular visiting instructors.

A third approach involves incentives to schools and students themselves instead of relying on institutions of higher education as providers of instruction. This "restoration strategy" recognizes that students from too many high schools (and minority students, in particular) have been slowed down by a variety of structural and environmental factors, and arrive at the college application line in the fall semester of grade 12 with 10 grades worth of work in their portfolios. If *all* our 4-year colleges adopted a rolling admissions cycle that takes in 11 percent of the target pool each month from November through July (yes, July!), students would be able (a) to recapture as much as a full year's worth of learning, and (b) to prepare for testing as late as April or May. Since test scores follow learning (not the other way around), disadvantaged students would benefit more from this recapturing of learning time. The so-called "bridge programs" that take place during the summer following high school graduation may be a helpful reinforcement in the restoration strategy, but would be more effective if they began two summers earlier (after the 10^{th} grade), on a much larger scale, and with follow-up cooperative curriculum-fortifying activities in the school district.

For those who doubt that opportunity to fill one's high school portfolio with an academic intensity and quality that would place one toward the top of the scale we constructed for ACRES and hence close the race gap in degree completion, I offer table 40. The data are a stunning endorsement of what curricular intensity and quality can do for African-American and Latino students in particular.

Some will argue that the Latino data in table 40 are too unstable (the standard errors are large, rendering statistical comparisons with the other race/ethnicity groups tenuous) and not representative of the Latino population in higher education. After all, for the High School & Beyond/Sophomores, 54 percent of the Latino students who continued their education after high school began in community colleges, a ratio far larger than that for any other race/ethnicity group[47]. That proportion, however, declined from 57 percent for the generation of the NLS-72, and has dropped to 49 percent for the NELS-88 cohort (the high school graduating class of 1992). This trend suggests that when we are able to validate this entire analysis using the NELS-88 college transcripts, we will witness less volatility and stronger statistical relationships.

Of the three core variables in table 40, only the curriculum variable is criterion-referenced. That is, the highest 40 percent does *not* necessarily describe relative position, particularly when the HIGHMATH criterion of trigonometry or higher is added. To repeat: in a happy paradox, *everybody* can be in the "highest 40 percent" on the curriculum measure. This is not the case for a test score measure, not matter what combination of tests we use. And it certainly is not the case for class rank, which by definition, is a relative measure.

Table 40.—Bachelor's degree completion rates for students in the top two quintiles of each component of ACRES, who entered 4-year colleges directly from on-time high school graduation, by race, High School & Beyond/Sophomore cohort, 1982-1993

	White	Black	Latino	Asian	All
ALL	75.4%	45.1%	60.8%	86.9%	72.1%
	(1.16)	(3.14)	(7.27)	(2.79)	(1.07)
Curriculum: Highest 40% and HIGHMATH beyond Algebra 2	85.7	72.6	79.3	89.0	84.8
	(1.44)	(4.98)	(7.34)	(3.47)	(1.33)
Test Scores Highest 40% of Combined Scale	80.5	67.1	66.6	94.7	79.9
	(1.17)	(3.66)	(8.38)	(1.90)	(1.09)
Class Rank/GPA Highest 40% of Combined Variable	78.9	58.8	57.0	84.9	77.1
	(1.26)	(4.56)	(7.44)	(2.95)	(1.19)

Notes: (1) Universe for "ALL" consists of all on-time high school graduates who entered 4-year colleges directly from high school, and whose college transcript files are not incomplete (Weighted N=859K); the universe for the three component groups adds high school records with positive values for all three components (Weighted N=805K). (2) standard errors are in parentheses. SOURCE: National Center for Education Statistics: High School & Beyond/ Sophomore cohort, NCES CD#98-135.

The Trouble With the "X-Percent" Solution

And yet it is to high school class rank and GPA that policy makers have turned for cheap and easy "solutions" in admissions to public institutions that exercise any degree of selectivity.

The policies are expressed in "take the top-X percent by class rank" formulas, even if, nationally, 19 percent of our secondary schools no longer use class rank, and 53 percent include non-academic courses in the calculation of GPA (College Board, 1998). This approach clearly does not acknowledge the *student* and his/her goals of completing a degree. As table 41 (an abbreviation of the data in table 40) makes abundantly clear, using class rank/GPA does the least for everybody, and actually *would have had a negative impact on Latinos* in the High School & Beyond/Sophomore cohort.

If we are genuinely interested in improving the degree completion rates of minority students, which of these indicators would we rather use? The answer, as they say, is a "no-brainer": the only field on which we can exercise change. A test score is a snapshot of performance on a Saturday morning. Secondary school grades—and the relative standing that they produce in "classes" where the student body may be constantly changing—carry as much reliability as a pair of dice (Elliott and Strenta, 1988). But the intensity and quality of curriculum is a cumulative investment of years of effort—by schools, teachers, and students, and provides momentum into higher education and beyond. It obviously pays off. The effects of grades and tests diminish in time, but the stuff of learning does not go away.

Table 41.—Comparative improvement in bachelor's degree attainment rates by moving into the top 40 percent on each component measure of the ACRES index, by race, High School & Beyond/Sophomore cohort, 1982-1993

Component:	White	Black	Latino	All
Curriculum	+10.4%	+27.5%	+18.5%	+12.7%
Test Scores	+ 5.1	+22.0	+ 5.8	+ 7.8
Class Rank/GPA	+ 3.5	+13.7	- 3.8	+ 5.0

SOURCE: National Center for Education Statistics: High School & Beyond/ Sophomore cohort, NCES CD#98-135.

Contemporary policy with respect to admissions at selective or highly selective public institutions, of course, does *not* focus on the top 40 percent of class rank/GPA (let along any other measure), rather on the top 10 percent (Texas) or 4 percent (California). The portrait in tables 40 and 41, to be fair, may not be the portrait resulting from a more restrictive threshold such as 10 percent or 4 percent. And a contemporary population may be very different from that of the HS&B/So. All this, of course, is speculative. The admissions line is not the commencement line. In an old saw, only time will tell. The common-sense odds, however, say that unless students in the "top X percent" by class rank/GPA *also* have the curriculum that comes from opportunity-to-learn, we may not be doing right by them.

Post-Matriculation Tools

The analysis of student progress in an age of multi-institutional attendance clearly advises us to keep the tool box handy. Access (entry) to higher education is not the dependent variable for students. Nor is mere retention to something called "year 2." From the multivariate analysis we learned just how large a role continuous enrollment plays in the degree completion. We know that keeping students enrolled, even for one course a term, is critical. How do we do that when students are highly mobile, when they behave like consumers, and when their loyalties to particular institutions are weak?

We possess the technology to answer that question in action. The dean's offices of this world have to know when a student intends to leave the institution, and must do everything in their persuasive power to ensure either that the student is transferring to another institution without a pause in the course of study, or that the student is connected to a course, anywhere, so that a potential break from academic momentum does not lengthen to irretrievable dimensions. In following the 16 percent of beginning 4-year college students in the BPS90 study who left higher education by the end of their first year (1989-1990), Horn (1998) found that 36 percent of this group returned within a year, but as the gap lengthened, the return rate fell. In replicating Horn's analysis with a longer time fame (1982-1993) and the broader population that was subject to our multivariate analysis, I found that the 26 percent of stop-outs who returned within a year completed degrees (associate's or bachelor's) at a 50 percent rate, but that any further delay in returning to college cut that completion rate in half.

Our other options include finding the student an on-line course from the increasing number of providers of Internet-based instruction, and not fretting if the provider is *not* a traditional 2-year or 4-year college. If the student does not own a computer and lives in the institution's immediate area, loan a computer to the student. Depending on student interests, occupational as well as academic, there are variations on this theme. If we are serious about helping students complete degrees, we can be creative. But in these situations, subsequent vigilant contact with students—by e-mail, by phone, by whatever—by advisors is necessary.

We found that a high DWI (Drops-Withdrawals-Incompletes) index worked significantly against degree completion. The situation could be ameliorated by institutional policies that both restrict the extent of withdrawals, incompletes, and no-credit repeats *and* play closer attention, in advisement, to student credit loads and time commitments. In others words, a restrictive policy alone does not help students unless it is accompanied by advisement actions that enable a student to complete a reasonable course load successfully.

This tool box is placed best in the offices of counselors and advisers, on the desks of those journalists and editorial writers who interpret higher education for the broad public, and, most of all, in the minds of students' friends and family. The tools derive from the principal features of the story-line. That story, brought to us by the wisdom of the U.S. Department of Education in establishing and maintaining its longitudinal studies, is a legacy from one generation to the next.

1. National Center for Education Statistics, *High School & Beyond Sophomore Cohort: 1980-92, Postsecondary Transcripts* (NCES 98-135). The CD includes not only the postsecondary transcript files, but also the high school transcript files, approximately 200 student-level variables constructed from the survey data, a labor market experience file, and an institutional file. The data from these selected files, while sufficient for most analyses of the life course histories of this cohort, can be merged with thousands of other variables on the original (1995) version of the HS&B/So restricted data set.

2. Surveys of the group were taken in 1980, 1982, 1984, 1986, and 1992. Postsecondary transcripts were gathered in 1993. Only four percent of the cohort was enrolled in postsecondary education in 1993, so for most students in the sample, the history ends at age 28/29 in 1992. A small number of students in the sample died between 1980 and 1992. The analysis file used in this study excludes those who passed away prior to 1984 on the grounds that, in terms of educational histories, they did not have the chance to complete degrees.

3. The BPS89-94 is an "event cohort" study, not an age cohort study. The initial group consisted of a national sample of people who were true first-time postsecondary students in the academic year, 1989-1990. These students ranged in age from 16 to over 50. The data collected for this group included neither high school nor college transcripts.

4. For example, during the five years of the Beginning Postsecondary Students Study (1989-1994), 18 percent of participants moved from dependent to independent status, 19 percent experienced a change in marital status, and 9 percent added children to their household (14 percent already had children when they started postsecondary education in 1989). Source: Data Analysis System, BPS90.

5. In most institutions of higher education, and certainly at a state university, the 3rd semester of calculus assumes that the student has previously studied elementary functions and analytic geometry. If one is placed in the 5th semester of college-level Russian, one can assume, at a minimum, prior study at the 4th semester level.

6. See, for example, Fernand Braudel's essay, "History and the Social Sciences," in Braudel (1980), pp. 25-54.

7. Of the HS&B/So students who took more than two remedial courses in college, the Census division x urbanicity of high school cells in which one finds the largest proportions (and in relation to their share of the origins of all postsecondary students) were:

Division x Urbanicity	Proportion of Remedial Students	Proportion of All Students
South Atlantic, suburban	10.4%	6.8%
East North Central, suburban	9.7	12.2
Pacific, suburban	9.6	8.3
Mid-Atlantic, suburban	6.8	9.8
Mid-Atlantic, urban	5.9	3.8
East North Central, urban	5.4	3.9
South Atlantic, rural	5.1	4.8
West South Central, rural	4.5	3.1

The lessons of such detail are not only that suburban high schools can be significant producers of remedial students, but that specific Region x Urbanicity configurations are over-represented in the origins of remedial students: South Atlantic and Pacific suburban schools and East North Central and Mid-Atlantic urban schools.

8. Given the 1,000+ course taxonomy in *The New College Course Map and Transcript Files* (U.S. Department of Education, 1995), the following illustrates what is/is not included in the aggregate category of "remedial courses." The examples are generalized. Institutional credit and grading policies help determine what is "remedial" in any given case. If, for example, in institution X, "Grammar and Usage" is indicated as a non-additive credit course with a non-standard grade (e.g. "Y"), it would be classified as remedial. If the same title appeared on a student record as a junior year course for someone who had previously taken courses in Shakespeare and creative writing and the credits were additive and the grades standard, the course would have been classified under linguistics. There were over 300,000 course entries in the HS&B/So transcript sample. Every entry was examined in this manner.

Included in the "Remedial" Aggregate	***Not* Included**
Basic Skills: Student Development	Student Orientation
Basic Academic Skills	Library Skills/Orientation/Methods
Remedial English; Developmental English, Punctuation, Spelling, Grammar, Basic Language Skills, Grammar and Usage	Reading & Composition, Exposition
Basic Writing, Writing Skills	Academic Writing, Informational Writing
Remedial/Basic Speech, Basic Oral Communication, Listening Skills	Fundamentals of Speech, Speech Communication, Effective Speech
Basic Reading, Reading Skills, Reading Comprehension	Speed Reading, Reading & Composition

Included in the "Remedial" Aggregate	***Not* Included**
Business Math: Pre-College, Business Business Arithmetic, Business Computations, Consumer Math	Math for Business/Econ, Math for Finance, Business Algebra
Arithmetic	Number Systems/Structures
Pre-College Algebra	Algebra for Teachers

9. We can drive home the point even further by comparing student accounts from BPS to transcript accounts from the HS&B/So in the matter of the type of remedial courses taken:

	BPS (student)	HS&B/So (transcript)
Any Remedial Mathematics	8.6%	33.7%
Any Remedial Reading	7.4	11.2

Hypothesis: students are more likely to know that they are in remedial reading than in pre-collegiate level mathematics.

10. To cite all the major studies that collapse high school curriculum in this manner would consume a dozen pages.

11. The procedure involves, first, matching all existing cases of students who show both SAT/ACT and senior test scores on their records, and determining the percentile on the senior test score that matched the *median* score on the SAT. It is not surprising that this percentile (54th) is higher than the mean for the senior test since the SAT/ACT test-taking population has been filtered by college-going intentions. The second step in imputing senior test score percentile from SAT and ACT scores is to call the median for each test the 54th percentile and to distribute the rest of the SAT/ACT scores in terms of percentiles.

12. The average was based on grades in non-remedial mathematics courses, English, all science courses, foreign language, history, and social studies. Grades in fine and performing arts and vocational courses were not included.

13. Faced with a similar situation, and because they had a much smaller sample and wished to reduce the number of missing cases to a minimum, Alexander and Eckland (1973) used a regression weight in the opposite direction: from the school principal's report of the student's class rank, in quintiles, to the student's self-reported "grade averages."

14. The five variables involve different combinations of Carnegie units in high school subject matter as follows:

	NWBASIC1	NWBASIC2	NWBASIC3	NWBASIC4	NWBASIC5
English	4.0	---	---	---	---
Mathematics	3.0	3.0	3.0	3.0	2.0
Science	3.0	3.0	3.0	3.0	2.0
Social St.	3.0	3.0	3.0	3.0	3.0
For. Langs.	2.0	---	2.0	---	---
Comput Sci.	0.5	0.5	---	---	---

For four of these variables, we have no indication of how much English is involved; and for three of them, no indication of how much foreign languages or computer science may be involved. Taking these constructions at face value, there is no hierarchical difference between NWBASIC2 and NWBASIC3. In fact, while one can set NWBASIC1 aside as an ideal, the only claim to a hierarchy is implicit in the numbering of the combinations. Those interested in the frequency counts for these variables can view them in the public release Data Analysis System (DAS) for High School & Beyond/Sophomore cohort included on the National Center for Education Statistics' CD#98-074.

15. The categories of mathematics available on a high school transcript sample from the late 1960s were: Algebra 1, Algebra 2, geometry, trig, calculus, general math (1,2,3, and 4), applied math (1 and 2), advanced math, and math not elsewhere classified (see Pallas and Alexander, 1983, p. 175). This is a very difficult list to configure in a categorical variable with intervals that clearly delineate a hierarchy. By the time the HS&B/So cohort was in high school a decade later, pre-calculus was a standard high school offering, pre-algebra courses were clearly identified, and statistics was specified, i.e. there was a great deal more specificity on the HS&B/So high school transcripts, and one could construct an HIGHMATH variable with the following values: Calculus, Pre-Calculus, Trigonometry, Algebra 2, Geometry, Algebra 1, Pre-Algebra/General Math 1 and 2/Arithmetic, and Indeterminable.

16. The early drafts of this monograph included a separate section in which the Academic Resources construction was replicated and tested using a newly-edited version of the NLS-72 high school records, and in which Altonji's work was described in more detail. At the advice of reviewers, this section was set aside for separate publication.

17. There are two versions of the high school transcripts in the HS&B/So data file, each based on a slightly different coding system. The HSTS version, on which this study relies, did not include an accounting for remedial English." The CTI version of the transcripts was merged for this variable. Fractional credits (less than 0.5) that were labeled "remedial English" in the CTI version of the transcripts were not deemed remedial since the "courses" at issue included tutorials and workshops, and these do not necessarily mean developmental work.

18. A criterion of 0.5 or more credits of computer science, as a dummy variable, was added at eight points along the scale to disaggregate lumps in the distribution. And at four points along the scale, credits earned in mathematics or science were added for the same purpose.

19. Pallas and Alexander (1983) attributed about 60 percent of the gap between men's and women's scores on the SAT-Q "to the sparse quantitative programs of study typically pursued

by girls in high school" (p. 181). The High School & Beyond/Sophomore data, however, do not show much of a divergence in the highest level of mathematics studied by men and women in high school, nor a significant divergence in composite senior year test scores *by* highest level of mathematics:

	Proportion Who Reached This Level of Mathematics in High School		Proportion Scoring in the Highest Quintile of the Senior Test Composite	
	Men	Women	Men	Women
Calculus	5.3	4.1	82.3	81.2
Pre-Calculus	4.7	3.9	66.5	63.7
Trigonometry	9.0	7.8	51.9	48.3
Algebra 2	21.3	23.9	31.0*	26.9*
<Algebra 2	59.7	60.3	6.9	5.8

Notes: (1) The universe consists of all HS&B/So high school graduates for whom highest level of mathematics could be determined and for whom a composite test score quintile was available. Weighted N=2.9M. (2) * $p \leq .05$.

20. Alexander, Riordan, Fennessey and Pallas define "high academic resources" as one standard deviation above the mean on ability (the senior test) and class rank, both within a college preparatory high school curriculum (p. 325). The principal problem with this formulation lies in class rank, since the variable is computed within-school, whereas the other components are not.

21. Using the 1994 survey of the NELS-88 cohort, i.e. two years after scheduled high school graduation, Berkner and Chavez (1997) first examined the pre-collegiate records of all students who said they had attended a 4-year college as of that date. They then judged *all* NELS-88 students with reference to the profiles of those who had entered 4-year colleges. Students were judged to be "college qualified" if their records evidenced at least one value on any of five criteria that would place them among the top 75 percent of 4-year college students for that criterion. The minimum values for "qualified" were: a class rank of the 46th percentile, an academic GPA of 2.7, an SAT combined score of 820, an ACT composite score of 19, or a NELS-88 test score (roughly the same test as used for the HS&B/So) of the 56th percentile. Curriculum in the form of the NWBASIC1 variable (see p. 13 above) was used to adjust degrees of "qualification," that is, it played a secondary role despite data that show it to be of primary importance. ACRES, in contrast, does not judge students with reference to isolated criteria, rather provides an analytic indicator of the general level of academic resource development toward which students can reach.

22. The most simplistic line of these inquiries use data from the Current Population Surveys of the Census Bureau despite the ambiguities in the way Census asks questions about "college" enrollment and (until 1992) attainment. The ambiguity produces extraordinarily volatile year-to-year enrollment rates, particularly by race, though analysts usually ignore the volatility. Other problems with time series in Census data involve the 1992 division of the question

concerning secondary school completion into two categories (diploma and equivalency) and the fact that the Current Population Surveys do not contain information on immigration status (Census focuses on college enrollment for the non-institutionalized civilian population age 18-24, and this group includes people who attended primary and/or secondary school in other countries). There are simply too many cross-currents and too much imputation in the data collection methodology of Census to rely on this source for precise estimates of college access or even degree completion (Pelavin and Kane, 1990). Other benchmark data assembled by the American College Testing Service, for example, come far closer to the longitudinal studies estimates and evidence far less volatility (see *Digest of Education Statistics, 1997*, tables 183 and 184, pp. 194-195).

23. The NELS-88 longitudinal study reminds those who tend to forget the importance of this factor in initial choice of college: 71 percent of high school seniors in 1992 cited location as a primary factor in choice. Source: Data Analysis System (DAS), National Education Longitudinal Study of 1988.

24. This was true particularly for those of Hispanic background. The cultural tone of persistence decisions in this population is very difficult to model, but critical to acknowledge. None of the national longitudinal studies account for the role of family and significant others after initial access to higher education. There is no reason to believe that a student will offer anything but an honest assessment to the true/false statement, "my family encourages me to continue attending this institution" (Cabrera, Nora, and Castaneda, 1993), but the relative power of the attitudes behind the statement may be very different for students in community colleges compared with those in 4-year colleges compared with those attending more than one institution.

25. McCormick's study excluded students who attended 4-year colleges but began their careers in other types of institutions (26 percent of all 4-year students and 20 percent of the bachelor's degree recipients in the HS&B/So), as well as 4-year college students whose 12th grade educational aspirations were less than a bachelor's degree (29 percent). McCormick also excluded credits earned by examination, credit equivalents of clock-hour courses, credits earned at less than 2-year schools, and credits earned before high school graduation. Some of these exclusions are unfortunate, but they should not detract from an instructive exposition.

26. Confining their interest only to completion at the first institution of attendance, the purpose of Astin, Tsui and Avalos' study was to demonstrate the difference between predicted and actual institutional graduation rates. The model is worth visiting. Starting with self-reported high school grades, and with stepwise feeding of SAT scores, gender and race, these authors found the adjusted R^2s increased from .281 to .325 in predicting 9-year completion rates within an institution (with gender adding a small amount to the R^2 and race adding almost nothing). What does that mean and how do we judge the results? An R^2 of .325 means that the model accounts for about one-third of the variance in what happened to this population. Given all the intervening behaviors of a 9-year period, one-third of the variance is a very strong number. As we will see, however, models that transcend the boundaries of a single institution are both more persuasive and produce even stronger estimates.

27. However, the proportion varies widely by type of true first institution of attendance and by combinations of institutions attended. For example, 24 percent of those whose first institution of attendance was a community college indicated they had been part-time students versus 9.8 percent of those who first entered comprehensive colleges. Among those who attended only 4-year colleges, 6.5 percent indicated part-time status at some time during their undergraduate careers, versus 20 percent of those engaged in alternating or simultaneous enrollment in 4-year and 2-year colleges.

28. Source: Data Analysis System (DAS), BPS90.

29. Unfortunately, Berkner, Cuccaro-Alamin, and McCormick did not make use of one of the most important filtering variables in the BPS90, namely the question of whether the respondent categorized themselves as a student who happens to be employed (63.7%) or an employee who happens to be a student (36.3%) The distinction ripples through the entire dataset and any interpretation of student careers. Here are some examples of how the differences in primary status played out in the first institution of attendance (1989-1990):

	Percent Primarily Students	Percent Primarily Employees
ALL	63.7	36.3
Level of First Institution		
4-year	77.5	22.5
2-year	49.1*	50.9*
<2-Year	38.0	62.0
Degree Working Toward		
None	23.9	76.1
Certificate	37.4	62.3
Associate's	53.8	46.2
Bachelor's	76.2	23.8
Enrollment Intensity		
Full-Time	73.6	26.2
Part-Time	31.6	62.0

Source: Data Analysis System (DAS), BPS90. * Not a statistically significant difference.

30. Hearn (1992) also warns us "to be suspicious of the measurement properties [validity, reliability] of aspirations, plans, and expectations indicators when the data are from the responses of middle or late adolescents." (p. 661).

31. Morgan (1996) translated the categorical variable of educational aspirations into years of schooling (e.g. a master's degree was worth 18 years, 6 of which were postsecondary). His notion was that if there were still differences in educational expectations of sub-groups after controlling for SES (which he standardized to a scale in which the mean=0 and SD=1 in order to merge the HS&B/So and NELS-88 cohorts), then we have to look elsewhere to explain the residual. With this methodology, Morgan found that, net of drop-outs, 29 percent of the students increased their expectations and 26 percent lowered their expectations between grades 10 and 12. This is a much less dramatic change than that revealed by the minimum educational level satisfaction questions treated as categorical variables.

32. The family income variable in the HS&B/So (as well as most other national data sets) is equally as tenuous as parental levels of education, but may be more important in analyses of college going, persistence, and completion. I chose to use the family income file for the HS&B/So base year (1980) prepared as a by-product of a report to NCES estimating families' capacity to finance higher education for their children (Dresch, Stowe, and Waldenberg, 1985). The "Dresch file" examined all attendant features of family and student, and removed outlying cases. The analytic problem with the "Dresch file" is that it reports income in eight unequal bands. The distribution appears reasonable, but there are no means here, no standard deviations, and no regular intervals on a continuous scale. When set on a grid against SES quintiles, we lose nearly 25 percent of the HS&B/So universe. As Sewell and Hauser (1975) effectively demonstrated, non-economic aspects of stratification are more important than the economic. So one should stick with the composite rather than lose such a large proportion of the sample.

33. NELS-88 provides both student and parent accounts, but 15 percent of the parents did not provide information on their highest levels of education, and 23 percent skipped the question on family income.

34. There were only 414 cases out of 14,825 in the data base that could be edited in this manner. In weighted numbers, the results indicate that we have been *over*-estimating first-generation college status by a *minimum* of 4 percent.

35. For students of "typical college age" in the Beginning Postsecondary Students Study of 1989-1994, we find the following with respect to factors in choosing the first institution of attendance, by general type of institution the student actually entered:

	4-Year College	2-Year College	<2-Year
School Close to Home An Important Factor in Choice	59.4%	75.4%	50.1%
First Institution Was 50 or Fewer Miles from Home	45.7%	90.0%	78.0%

Source: National Center for Education Statistics: Beginning Postsecondary Students, 1989-1994, Data Analysis System.

36. Transcript practices with respect to study abroad are highly variable, and require hand-and-eye reading to identify. Some may be highly explicit, e.g. "University of Heidelberg" or "Monterrey Semester." Others may use an abbreviation, e.g. "Bensacon," and the transcript reader simply has to know that a major French language training center is being referenced. In still other cases, the sequence of courses for a history major shifts in the spring term to "Art and Architecture of Florence," "The Age of Machiavelli," "Advanced Italian Conversation," *Il Paradiso*, and "Antonioni and Italian Cinematic Realism." The reader knows that it is highly unlikely that the student is attending school in the United States.

37. How does this estimate stack up against current (1999) claims of massive numbers of bachelor's degree students doing post-baccalaureate work *for credit* in community colleges? No one has conducted a complete census, let alone a census with unduplicated headcounts. But let us speculate. The HS&B/So is only one high school graduating class. Assume that there are ten high school graduating classes "in play" at the present moment, and that in each class of eventual bachelor's degree recipients there has been a 10 percent increase in the proportion attending a community college after the BA. This fairly generous set of assumptions results in an estimate of about a half-million credit students, or 9 percent of community college students currently enrolled for credit.

38. The outcome effects of selectivity have been remarkably consistent for two generations of college graduates: both the NLS-72 and HS&B/So:

Relation of selectivity of bachelor's degree-granting institution to graduate school attendance and GPA in two cohorts of college graduates, 1972-1993

	NLS-72 1972-84	HS&B/Sophomores 1982-93	
Graduate School Attendance by Age 30			
Highly Selective	39.7%	43.0%	
Selective	34.6	31.0	
Non-Selective	23.1	17.9	
Mean (S.D.) Undergraduate GPA			
Highly Selective	3.16 (.51)	3.13 (.42)	Effect size=.07
Selective	3.01 (.50)	2.96 (.45)	Effect size=.11
Non-Selective	2.92 (.48)	2.86 (.42)	Effect size=.13

NOTES: (1) The universes are confined to students who earned bachelor's degrees and for whom an undergraduate GPA could be determined. NLS-72 weighted N=732k; HS&B/So weighted N=935k. (2) Differences in estimates of graduate school attendance rates are significant at $p \leq .05$. (3) Effect sizes for changes in mean GPA indicate no change. SOURCE: NCES: (1) National Longitudinal Study of the High School Class of 1972; (2) High School & Beyond/Sophomore Cohort, NCES CD#98-135.

39. The proportion of all undergraduate grades that were drops, no-penalty withdrawals, and no-penalty incompletes rose from 4 percent for the NLS-72 cohort to over 7 percent for the HS&B/So.

40. The variable, HSBSTAT, on the 1998 restricted CD release dataset, divides the entire universe of HS&B/So students into five groups with reference to their postsecondary status. Of these groups, only three are in the potential universe for analysis: (1) students for whom postsecondary transcripts were received (8215, of whom 14 had died by 1986 and are not included), (2) students for whom transcripts were received, but the content of the transcripts was entirely GED-level or basic skills work (180, of which very few had attended 4-year colleges), and (3) students for whom transcripts were requested but not received, yet for whom the evidence allows imputation of college attendance (478). In the basic universe for multivariate analyses, the expansion group comes from the third of these categories. Relying on survey data, it was possible to determine the *order* of institutional attendance for students in the third of these groups, hence values for variables such as Transfer and No-Return.

41. Of all students who entered postsecondary education, 7 percent of the men and 16.2 percent of the women became parents by 1986. Among students who attended 4-year colleges at any time, the figures were 3.3 percent (men) and 7.5 percent (women).

42. Had I not edited the "Children" variable to exclude contradictory and out-of-scope information, its contribution to the explanatory power of the model, as well as the overall explanatory power of the model, would have been 0.2 percent higher. What this tells us, albeit very indirectly, is that the people who provided contradictory or out-of-scope information about having children probably did not complete bachelor's degrees.

43. The statistical software packages commonly used for regression models (SAS, SPSS, STATA, and others) first set up all the variables in a correlation matrix. Only those independent variables whose correlations with the dependent variable (in our case, bachelor's degree completion) are statistically significant at $p \leq .05$ are allowed into the regression equation. $p \leq .05$ is a default value. One can change this criterion to be more or less generous. I have chosen to be more generous throughout the models in this study by setting the selection criterion to $p \leq .20$. The reader will note that, under those conditions, variables of marginal significance will not hold up when they are asked to take on an explanatory role.

44. The California State University at Fullerton is *not* the California State University at Monterey Bay. Queens College of the City University of New York is *not* the John Jay College of Criminal Justice (also part of CUNY). Even a multi-campus institution such as Southern Illinois University presents different environments for students who move among the campuses.

45. For the academic satisfaction index, responses to questions concerning the quality of teachers, the quality of instruction, curriculum, intellectual life of the school, and personal intellectual growth were combined. For the environmental index, responses to questions concerning the social life, cultural activities, recreation facilities, and adequacy of services and other facilities were combined.

46. Cabrera, Nora, and Castaneda (1993) ran into similar difficulties with a satisfaction measure of college finances, although the question on their survey involved total financial support ("I am satisfied with the amount of financial support [grants, loans, family assistance, and jobs] I have received while attending . . .") and not costs. In an integrated persistence model, they found that the effects of this type of satisfaction were indirect, and expressed through academic integration and GPA. Even then, the structural coefficients of these effects were rather weak (.138 for academic integration and .104 for GPA).

47. White, African-American, and Asian-American students have very similar enrollment rates in community colleges, no matter how those rates are represented. In the following table, with three "parsings," it is obvious how much Latino attendance patterns differ from the others, and how much the community college means to Latinos communities:

Proportions of students attending community colleges under three different notions of "attendance," by race/ethnicity, High School & Beyond/Sophomores, 1982-93

	First Institution Attended	Ever Attended	The *Only* Institution Attended
White	39%	50%	25%
Black	39	51	24
Asian	36	55	20
Latino	54	66	40

Note: Comparisons of Latino estimates with others are significant at $p \leq .05$; comparisons among other race/ethnicity groups are not significant. Source: National Center for Education Statistics: High School & Beyond/Sophomore cohort. NCES CD #98-135.

REFERENCES

Adelman, C. 1983. *Devaluation, Diffusion and the College Connection: a Study of High School Transcripts, 1964-1981*. Report prepared for the National Commission on Excellence in Education. Washington, DC: U.S. Department of Education. ERIC document #228-244.

Adelman, C. 1994. *Lessons of a Generation: Education and Work in the Lives of the High School Class of 1972*. San Francisco: Jossey-Bass.

Adelman, C. 1995. *The New College Course Map and Transcript Files: Changes in Course-Taking and Achievement, 1972-1993*. Washington, DC: U.S. Department of Education.

Adelman, C. 1996. "The Thin Yellow Line: Editing and Imputation in a World of Third Party Artifacts." In Alvey, W. and B. Jamerson (eds.), *Data Editing Exposition*. Washington, DC: Office of Management and Budget, pp. 173-181.

Adelman, C. 1998. *Women and Men of the Engineering Path: a Model for Analyses of Undergraduate Careers*. Washington, DC: U.S. Department of Education and the National Institute for Science Education.

Adelman, C. 1999. "Crosscurrents and Riptides: Asking About the Capacity of the Higher Education System." *Change*, vol. 31. no. 1, pp. 20-27.

Akerhielm, K., J. Berger, M. Hooker, and D. Wise. 1998. *Factors Related to College Enrollment*. Princeton, NJ: Mathtech, Inc. [Final Report to the Office of Planning and Evaluation Services, U.S. Department of Education, contract #EA94078001].

Alexander, K.L. and Cook, M.A. 1979. "The Motivational Relevance of Educational Plans: Questioning the Conventional Wisdom." *Social Psychology Quarterly*, vol. 42, no. 3, pp. 202-213.

Alexander, K.L. and Cook, M.A. 1982. "Curricula and Coursework: Surprise Ending to a Familiar Story." *American Sociological Review*, vol. 47, pp. 626-640.

Alexander, K. L. and B.K. Eckland. 1973. *Effects of Education on the Social Mobility of High School Sophomores Fifteen Years Later (1955-1970)*. Chapel Hill, NC: Institute for Research in Social Science, University of North Carolina.

Alexander, K.L. and Eckland, B.K. 1977. "High School Context and College Selectivity: Institutional Constraints in Educational Stratification." *Social Forces*, vol. 56, pp. 166-188.

Alexander, K.L., Holupka, S., and Pallas, A.M. 1987. "Social Background and Academic Determinants of Two-Year versus Four-Year College Attendance: Evidence from Two Cohorts a Decade Apart." *American Journal of Education*, vol. 96, pp. 56-80.

Alexander, K.L., McPartland, J.M. and Cook, M.A. 1981. "Using Standardized Test Performance in School Effects Research." *Research in Sociology of Education and Socialization*, vol. 2, pp. 1-33.

Alexander, K.L. and Pallas, A.M. 1984. "Curriculum Reform and School Performance: an Evaluation of the 'New Basics." *American Journal of Education*, vol. 92, pp. 391-420.

Alexander, K.L., Pallas, A.M. and Holupka, S. 1987. "Consistency and Change in Educational Stratification: Recent Trends Regarding Social Background and College Access." *Research in Social Stratification and Mobility*, vol. 6, pp. 161-185.

Alexander, K.L., Riordan, C., Fennessey, J., and Pallas, A.M. 1982. "Social Background, Academic Resources, and College Graduation: Recent Evidence from the National Longitudinal Survey." *American Journal of Education*, vol. 91, pp. 315-333.

Aldrich, J. H. and Nelson, F.D. 1984. *Linear Probability, Logit and Probit Models*. Beverly Hills, CA: Sage University Press.

Altonji, J.G. 1992. "The Effects of High School Curriculum on Education and Labor Market Outcomes." *Journal of Human Resources*, vol. 30, no. 3, pp. 409-438.

Astin, A.W. 1972. *College Dropouts: a National Profile*. Washington, DC: American Council on Education.

Astin, A.W. 1977. *Four Critical Years*. San Francisco: Jossey-Bass.

Astin, A.W. 1993. *What Matters in College? 'Four Critical Years' Revisited.* San Francisco: Jossey-Bass.

Astin, A.W., Hemond, M.K., and Richardson, G.T. 1982. *The American Freshman: National Norms for Fall 1982*. Los Angeles: Higher Education Research Institute, UCLA.

Astin, A.W., Tsui, L. and Avalos, J. 1996. *Degree Attainment Rates at American Colleges and Universities: Effects of Race, Gender, and Institutional Type*. Los Angeles: Higher Education Research Institute, UCLA.

Baker, T.L. and Velez, W. 1996. "Access to and Opportunity in Postsecondary Education in the United States: a Review." *Sociology of Education*. Extra Issue, pp. 82-101.

Bean, J.P. 1980. "Dropouts and Turnover: the Synthesis and Test of a Causal Model of Student Attrition." *Research in Higher Education*, vol. 12, pp. 155-187.

Bean, J.P. 1982. "Student Attrition, Intentions, and Confidence: Interaction Effects in a Path Model." *Research in Higher Education*, vol. 17, pp. 291-319.

Bean, J.P. 1983. "The Application of a Model of Turnover in Work Organizations to the Student Attrition Process." *Review of Higher Education*, vol. 6, pp. 129-148.

Bean, J.P. and Metzger, B.S. 1985. "A Conceptual Model of Non-traditional Undergraduate Student Attrition." *Review of Educational Research*, vol. 55, no. 4., pp. 485-540.

Berkner, L.K. and Chavez, L. 1997. *Access to Postsecondary Education for the 1992 High School Graduates*. Washington, DC: National Center for Education Statistics.

Berkner, L.K., Cuccaro-Alamin, S., and McCormick, A.C. 1996. *Descriptive Summary of 1989-90 Beginning Postsecondary Students: 5 Years Later.* Washington, DC: National Center for Education Statistics.

Blumberg, A.S., Lavin, D.E., Lerer, N, and Kovath, J. 1997. *Graduation Rates and Inferences of Institutional Effectiveness*. Paper presented at the Annual Forum of the Association for Institutional Research.

Braudel, F. 1980. *On History*. Chicago: University of Chicago Press.

Burbridge, L.C. 1991. *The Interaction of Race, Gender, and Socioeconomic Status in Educational Outcomes*. Wellesley, MA: Wellesley College Center for Research on Women.

Cabrera, A.F. 1994. "Logistic Regression Analysis in Higher Education: an Applied Perspective," in J.C. Smart (ed.), *Higher Education: Handbook of Theory and Research*, vol. X. New York: Agathon Press, pp. 225-256.

Cabrera, A.F. and Bernal, E.M. 1998. *Association Between SES and Ethnicity Across Three National Databases*. University Park, PA: Center for the Study of Higher Education.

Cabrera, A.F., Castaneda, M.B., Nora, A. and Hengstler, D. 1991. "The Convergence Between Two Theories of College Persistence." *Journal of Higher Education*, vol. 63, pp. 143-164.

Cabrera, A.F., Nora, A., and Castaneda, M.B., 1993. "College Persistence: Structural Equations Modeling Test of an Integrated Model of Student Retention." *Journal of Higher Education*, vol. 64, pp. 123-139.

Carroll, C.D. 1989. *College Persistence and Degree Attainment for 1980 High School Graduates: Hazards for Transfers, Stopouts, and Part-Timers*. Washington, DC: National Center for Education Statistics.

Cheney, B. and Farris, E. 1991. *Survey on Retention at Higher Education Institutions*. Washington, DC: National Science Foundation, National Endowment for the Humanities, and U.S. Department of Education. Higher Education Surveys Report. Survey #14.

Clubb, J.M., E.W. Austin, and G.W. Kirk, Jr. 1989. *The Process of Historical Inquiry: Everyday Lives of Working Americans*. New York: Columbia University Press.

The College Board. 1998. *High School Grading Policies*. Research Notes RN-04. New York City, NY: Author.

Connerton, P. 1989. *How Societies Remember*. Cambridge, UK: Cambridge Univ. Press.

Cuccaro-Alamin, S. and Choy, S.P. 1998. *Postsecondary Financing Strategies: How Undergraduates Combine Work, Borrowing, and Attendance*. Washington, DC: National Center for Education Statistics.

DesJardins, S.L., Ahlburg, D.A. and McCall, B.P. 1997. *Using Event History Methods to Model the Different Modes of Student Departure from College*. Paper presented at the Annual Forum of the Association for Institutional Research.

Dey, E.L. and Astin, A.W. 1993. "Statistical Alternatives for Studying College Student Retention: a Comparative Analysis of Logit, Probit, and Linear Regression." *Research in Higher Education*, vol. 34, no. 5, pp. 569-581.

Elliott, R. and Strenta, A.C., 1988. "Effects of Improving the Reliability of the GPA on Prediction General and on Comparative Predictions for Gender and Race Particularly." *Journal of Educational Measurement*, vol. 25, no. 4, pp. 333-347.

Fetters, W.B., Stowe, P.S. and Owings, J.A. 1984. *Quality of Responses of High School Students to Questionnaire Items*. Washington, DC: National Center for Education Statistics.

Flint, T.A. 1992. "Parental and Planning Influences on the Formation of Student College Choice Sets." *Research in Higher Education*, vol. 33, no. 6, pp. 689-708.

Garet, M.S. and DeLaney, B. 1988. "Students, Courses, and Stratification." *Sociology of Education*, vol. 61, no. 2, pp. 61-77.

Grandy, J. 1998. "Persistence in Science of High-Ability Minority Students." *Journal of Higher Education*, vol. 69, no. 6, pp. 589-620.

Haskins, L. and K. Jeffrey, 1990. *Understanding Quantitative History*. New York: McGraw-Hill.

Hauser, R.M. and Anderson, D.K. 1991. "Post-High School Plans and Aspirations of Black and White High School Seniors: 1976-86." *Sociology of Education*, vol. 64, pp. 22-30.

Hearn, J.C. 1984. "The Relative Roles of Academic, Ascribed, and Socioeconomic Characteristics in College Destination." *Sociology of Education*, vol. 57, pp. 22-30.

Hearn, J.C. 1988. "Determinants of Postsecondary Attendance: Some Implications of Alternative Specifications of Enrollment." *Educational Evaluation and Policy Analysis*, vol. 10, no. 2, pp. 172-185.

Hearn, J.C. 1991. "Academic and Non-Academic Influence on the College Destinations of 1980 High School Graduates." *Sociology of Education*, vol. 64, pp. 158-71.

Hearn, J.C. 1992. "Emerging Variations in Postsecondary Attendance Patterns: an Investigation of Part-time, Delayed, and Nondegree Enrollment." *Research in Higher Education*, vol. 33, pp. 657-87.

Horn, L. 1996. *Nontraditional Undergraduates*. Washington, DC: National Center for Education Statistics.

Horn, L. 1997. *Confronting the Odds: Students at Risk and the Pipeline to Higher Education*. Washington, DC: National Center for Education Statistics.

Horn, L. 1998. *Stopouts or Stayouts? Undergraduates Who Leave College in Their First Year*. Washington, DC: National Center for Education Statistics.

Jackson, G.A. 1988. "Did College Choice Change During the 1970s?" *Economics of Education Review*, vol. 7, no. 1, pp. 15-27.

Kanarek, E.A. 1989. *Exploring the Murky World of Admissions Predictions*. Paper presented at the 1989 National Forum of the Association for Institutional Research.

Kanoy, K.W., Wester, J., and Latta, 1989. "Predicting College Success of Freshman Students Using Traditional, Cognitive, and Psychological Measures." *Journal of Research and Development in Education*, vol. 22, pp. 65-70.

Karen, D. 1991. "Politics of Class, Race, and Gender: Access to Higher Education in the United States, 1960-86." *American Journal of Education*, vol. 99, pp. 208-37.

Kroc, R., Howard, R., Hull, P. and Woodard, D. 1997. *Graduation Rates: Do Students' Academic Program Choices Make a Difference?* Paper presented at the 1997 National Forum of the Association for Institutional Research.

Lang, M. 1992. "Barriers to Blacks' Educational Achievement in Higher Education: a Statistical and Conceptual Review." *Journal of Black Studies*, vol. 22, pp. 510-22.

Lee, V.E. and Frank, K.A. 1990. "Students' Characteristics That Facilitate the Transfer from Two-Year to Four-Year College." *Sociology of Education*, vol. 63, pp. 178-93.

Lee, V.E., Mackie-Lewis, C., and Marks, H.M. 1993. "Persistence to the Baccalaureate Degree for Students Who Transfer from Community College." *American Journal of Education*, vol. 102, pp. 80-114.

Lewis, L. and E. Farris. 1997. *Remedial Education at Higher Education Institutions in Fall 1995*. Washington, DC: National Center for Education Statistics.

Madigan, T. 1997. *Science Proficiency and Course Taking in High School*. Washington, DC: National Center for Education Statistics.

Marchese, T. 1998. "Not-So-Distant Competitors." *AAHE Bulletin*, vol. 15, no. 9. pp. 3-7.

Mare, R.D. 1980. "Social Background and School Continuation Decisions." *Journal of the American Statistical Association*, vol. 75, pp. 295-305.

McConnell-Castle, E. 1993. "Minority Student Attrition Research: Higher Education's Challenge for Human Resource Development." *Educational Researcher*, vol. 22, pp. 24-30.

McCormick, A.C. 1997. *Transfer Behavior Among Beginning Postsecondary Students: 1989-94*. Washington, DC: National Center for Education Statistics.

McCormick, A.C. 1999. *Credit Production and Progress Toward the Bachelor's Degree: an Analysis of Postsecondary Transcripts for Beginning Students at 4-Year Institutions*. Washington, DC: National Center for Education Statistics.

Medrich, E.A., Brown, C.L. and Henke, R.R. 1992. *Overview and Inventory of State Requirements for School Coursework and Attendance*. Washington, DC: National Center for Education Statistics.

Menard, S. 1995. *Applied Logistic Regression Analysis*. Thousand Oaks, CA: Sage Publications.

Monk, D.H. and Haller, E.J. 1993. "Predictors of High School Academic Course Offerings: the Role of School Size." *American Educational Research Journal*, vol. 30, pp. 3-21.

Morgan, S. L. 1996. "Trends in Black-White Differences in Educational Expectations: 1980-92." *Sociology of Education*, vol. 69, pp. 308-319.

Mow, S.L. and Nettles, M.T. 1990. "Minority Access to, and Persistence and Performance in College: a Review of the Trends and Research Literature." In J.C. Smart (ed.), *Higher Education: Handbook of Theory and Research* (Vol. 6, pp. 105-135). New York: Agathon Press.

National Center on the Educational Quality of the Workforce. 1998. "Understanding Employers' Perceptions of College Graduates." *Change*, vol 30., no. 3, pp. 47-50.

Nora, A. 1987. "Determinants of Retention Among Chicano College Students: a Structural Model." *Research in Higher Education*, vol. 26, pp. 31-59.

Nora, A. and Rendon, L.I. 1990. "Determinants of Predisposition to Transfer Among Community College Students: a Structural Model." *Research in Higher Education*, vol. 31, pp. 225-235.

Nora, A., Attinasi, L.C. and Matonak, A. 1990. "Testing Qualitative Indicators of Precollege Factors in Tinto's Attrition Model: a Community College Student Population." *Review of Higher Education*, vol. 13, pp. 337-356.

Orwell, G. 1954. "Politics and the English Language," in *A Collection of Essays by George Orwell*. New York: Doubleday & Co., pp. 162-176.

Pallas, A.M. and Alexander, K.L. 1983. "Sex Differences in SAT Performance: New Evidence of the Differential Course Work Hypothesis." *American Educational Research Journal*, vol. 20, no. 2, pp. 165-182.

Pascarella, E.T. and Terenzini, P.T. 1980. "Predicting Freshman Persistence and Voluntary Dropout Decisions from a Theoretical Model." *Journal of Higher Education*, vol. 51, pp. 60-75.

Pascarella, E.T. and P.T. Terenzini. 1981. *How College Affects Students*. San Francisco: Jossey-Bass.

Pascarella, E.T., Terenzini, P.T. and Wolfle, L.M. 1986. "Orientation to College and Freshman Year Persistence/Withdrawal Decisions." *Journal of Higher Education*, vol. 56, pp. 155-74.

Pedhazur, E.J. 1982. *Multiple Regression in Behavioral Research*. New York: Holt, Rinehart and Winston.

Pelavin, S.H. and M. Kane. 1990. *Changing the Odds: Factors Increasing Access to College*. New York City, NY: the College Board.

Peng, S.S. 1977. *Transfer Students Among Institutions of Higher Education*. Washington, DC: U.S. Department of Education.

Rock, D.A., Hilton, T.L., Pollock, J., Ekstrom, R.B., and Goertz, M.E. 1985. *Psychometric Analysis of the NLS and the High School and Beyond Test Batteries*. Washington, DC: National Center for Education Statistics.

Ronco, S. L. 1996. *How Enrollment Ends: Analyzing the Correlates of Student Graduation, Transfer and Dropout with a Competing Risks Model*. AIR Professional File, No. 61. Tallahassee, Fla.: Association for Institutional Research

Rosenbaum, J.E. 1976. *Making Inequality: the Hidden Curriculum of High School Tracking*. New York City, NY: John Wiley & Sons.

Rosenbaum, J.E. 1980. "Track Misperceptions and Frustrated College Plans." *Sociology of Education*, vol. 53, pp. 74-88.

Schroeder, L.D., D. L. Sjoquist, and P.E. Stephan. 1986. *Understanding Regression Analysis*. Newbury Park, Calif., Sage Publications.

Sewell, W.H. and R.M. Hauser. 1975. *Education, Occupation, and Earnings*. New York: Academic Press.

Skinner, C.J., Holt, D. and Smith, T.M.F. (eds.) 1989. *Analysis of Complex Surveys*. New York: John Wiley & Sons.

Smart, J. and Pascarella, E. 1987. "Influences on the Intent to Reenter Higher Education." *Journal of Higher Education*, vol. 58, pp. 306-322.

Smith, D.L. 1992. "Validity of Faculty Judgments of Student Performance: Relationship Between Grades and Credits Earned and External Criterion Measures." *Journal of Higher Education*, vol. 63, pp. 329-340.

Snyder, T. 1987. *Digest of Education Statistics, 1987.* Washington, DC: National Center for Education Statistics.

Spade, J.Z., Columba, L. and Vanfossen, B.E. 1997. "Tracking in Mathematics and Science: Courses and Course-Selection Procedures." *Sociology of Education*, vol. 70, pp. 108-127.

Spanard, J.A. 1990. "Beyond Intent: Reentering College to Complete the Degree." *Review of Educational Research*, vol. 60, pp. 309-44.

St. John, E.P. 1991. "What Really Influences Minority Attendance? Sequential Analyses of the High School and Beyond Sophomore Cohort." *Research in Higher Education*, vol. 32, pp. 141-158.

St. John, E.P. and Noell, J. 1989. "The Effects of Student Financial Aid on Access to Higher Education: an Analysis of Progress with Special Consideration of Minority Enrollment." *Research in Higher Education*, vol. 30, no. 6., pp. 563-582.

St. John, E.P. and Noell, J. 1990. "Price Response in Persistence Decisions: an Analysis of the High School and Beyond Senior Cohort." *Research in Higher Education*, vol. 31, no. 4, pp. 387-403.

St. John, E.P., Kirshstein, J.R., and Noell, J. 1991. "The Effects of Financial Aid on Persistence: a Sequential Analysis." *Review of Higher Education*, vol .14., no. 3, pp. 383-406.

Stampen, J. O. and Cabrera, A.F. 1986. "Exploring the Effects of Student Aid on Attrition." *Journal of Student Financial Aid*, vol. 16, pp. 28-40.

Tierney, W.G. 1992. "An Anthropological Analysis of Student Participation in College." *Journal of Higher Education*, vol. 63, pp. 603-17.

Tinto, V. 1975. "Dropout from Higher Education: a Theoretical Synthesis of Recent Research." *Review of Educational Research*, vol. 45, no. 1, pp. 89-125.

Tinto, V. 1982. "Limits of Theory and Practice in Student Attrition." *Journal of Higher Education*, vol. 53, pp. 687-700.

Tinto, V. 1993. *Leaving College: Rethinking the Causes and Cures of Student Attrition* (2nd ed.). Chicago: Univ. of Chicago Press.

Tracey, T.J. and Sedlacek, W.E. 1987. "Prediction of College Graduation Using Noncognitive Variables by Race." *Measurement and Evaluation in Guidance*, vol. 16, pp. 171-178.

Tuma, J., Gifford, A., Horn, L., and Hoachlander, E.g. 1989. *Enrollment Trends in Academic Education in American Public High Schools, 1969 to 1987.* Berkeley, CA: MPR Associates.

Velez, W. 1985. "Finishing College: the Effects of College Type." *Sociology of Education*, vol. 58, pp. 191-200.

Waite, L.J. and Moore, K.A. 1978. "The Impact of Early First Birth on Young Women's Educational Attainment." *Social Forces*, vol. 56, pp. 845-865.

Western Interstate Commission for Higher Education and the College Board. 1998. *Knocking at the College Door: Projections of High School Graduates by State and Race/Ethnicity, 1996-2012.* Boulder, CO and New York: Authors.

Appendix A: Technical Notes and Guidance

There are many tables in this document, both in the text and in the notes. Some are derived or constructed from other published sources or from the Data Analysis System (DAS) presentations of data sets published on CD-ROM by the National Center for Education Statistics. But most of the tables in this publication were prepared using special analyses files created from the High School and Beyond/Sophomore Cohort (HS&B/So) longitudinal study, and it is helpful to know something about the statistical standards that lie behind these tables and the decision rules that were used in presenting the data.

Weights and Adjusted Weights

The populations in all NCES age-cohort longitudinal studies are national probability samples first drawn when the students were in high school or middle school. In the case of the HS&B/So, the design involved first, a stratified sample of secondary schools with an over-sampling of schools in minority areas, and a random sampling of 10th grade students within those schools. The original sample was then weighted to match the national census of all 10th–graders in 1980 (about 3.7 million people). Each participant carries a weight in inverse proportion to the probability that he or she would be selected by chance. The HS&B/So base year sample was what statisticians call "robust": 28,000. After the base year, every subsequent survey was a subset of the original, and the weights carried by participants are modified accordingly. In the penultimate survey of the HS&B/So in 1992, there were 12,640 respondents out of 14,825 surveyed (of whom 155 had died). The postsecondary transcript file for the HS&B/So has 8,395 cases, and the special version of the transcript file used in this study has 8,873 cases. These are still very robust numbers. They represent populations in the millions. By the conclusion of any of these longitudinal studies, a student is carrying a half-dozen different weights, depending on what question is asked.

For the High School and Beyond cohort, for example, I used three basic weights in the tables in this study: a "senior year" weight for a question such as the relationship between the highest level of mathematics studied in high school and whether someone eventually earns a bachelor's degree; a "primary postsecondary transcript weight" for analyses of degree attainment for anyone for whom the evidence indicates attended a 4-year college; a "secondary transcript weight" for any question that would be compromised if students with incomplete postsecondary transcript records were included. Where correlation matrices and multivariate analyses are involved, each of these basic weights must be modified by the population that possesses positive values for all variables in an equation.

For example, the basic senior year weight for a group of students confined to those who graduated from high school on time and for whom we have SES data, known race, and all three components of ACRES (senior year test, class rank/GPA, and curriculum intensity & quality), would be modified as follows:

Weight2=senior year weight/(2320762/8844)

The numbers in parentheses are the weighted N and raw N derived from a simple cross-tabulation of any two of the variables in this set, using the senior year weight. Weight2 is what is carried into a correlation matrix for students with this set of variables. In the execution of the correlation matrix itself, I did not allow for pairwise missing values. The software program is instructed, "NOMISS." This decision derives from the historical approach to the data as discussed in the "Introduction" to this monograph.

Standard Errors and Significance Testing

More important are issues of standard errors of measurement and significance testing. What you see in the tables are estimates derived from samples. Two kinds of errors occur when samples are at issue: errors in sampling itself, particularly when relatively small sub-populations are involved, and non-sampling errors. Non-sampling errors are serious matters. Good examples would include non-response to specific questions in a survey or missing college transcripts. Weighting will not address the panoply of sources of non-sampling errors.

The effects of sampling and non-sampling errors ripple through data bases, and, to judge the accuracy of any analysis, one needs to know those effects. When the unit of analysis is the student, this is a straightforward issue. When we ask questions about combinations of institutions attended, bachelor's degree completion rates by selectivity of first institution of attendance, or highest level of mathematics studied in high school, we are asking questions about non-repetitive behaviors of people who were sampled. To judge comparisons in these cases we use the classic "Student's *t*" statistic that requires standard errors of the mean. But because the longitudinal studies were not based on simple random samples of students, the technique for generating standard errors involves a more complex approach known as the Taylor series method. For the descriptive statistics in this report, a proprietary program incorporating the Taylor series method, called STRATTAB, was used.

It is important to note that STRATTAB will provide neither estimates nor standard errors for any cell in a table in which the unweighted N is less than 30. For those cells, the program shows "LOW N." Table 19 on page 48 illustrates the frequency of LOW N cells that occur when one is making multiple comparisons among categories of an independent variable.

Most of the tables in this monograph include standard errors of the estimates and/or an indication of which comparisons in the table are significant at the $p \leq .05$ level using the classic "Student's *t*" test. The text often discusses these cases, and, when appropriate to the argument, offers the *t* statistic. A reader interested in comparing categories of a dependent variable that are not discussed can use the standard errors and employ the basic formula for computing the "Student's *t*":

$$t = (P_1 - P_2) / \sqrt{(se_1^2 + se_2^2)}$$

where P_1 and P_2 are the estimates to be compared and se_1 and se_2 are the corresponding standard errors. If, in this case, $t \geq 1.96$, you have a statistically significant difference such that the probability that this observation would occur by chance is less than 1:20. In the case of multiple comparisons, the critical value for t rises following the formula for Bonferroni Tests: if H comparisons are possible, the critical value for a two-sided test is $Z_{(1-.05/2H)}$. For a table showing the Z required to ensure that $p \leq .05/H$ for particular degrees of freedom, see Dunn, O.J., "Multiple Comparisons Among Means," *Journal of the American Statistical Association*, vol. 56 (1961), p. 55.

Design Effects and the Adjustment of Standard Errors

In multivariate analyses of a stratified sample such as any of the NCES longitudinal studies, it is necessary to adjust the standard errors produced with the software package (SPSS or SAS) by the average design effect, or DEFT (see Skinner, Holt, and Smith, 1989). Software packages such as SPSS or SAS assume simple random sampling when computing standard errors of parameter estimates. The DEFT for any population is computed by dividing the simple standard error by the Taylor Series error. The design effects for the HS&B/So populations used in the tables of this study range from 1.49 to 1.64.

These design effects are then carried over into the determination of significance in correlation matrices and regression analyses.

For correlation matrices, the formula for the standard two-tailed t-test thus becomes:

Step 1: Determine the adjusted standard error of the correlation coefficient (r).

$$\sqrt{\frac{(1-r)^2}{\div \; N-2}} \times \text{DEFT}$$

Step 2: Determine t = r / s.e.(r)

For regression analyses, the formula for the standard two-tailed t-test thus becomes:

Step 1: $$\left(\frac{B}{(s.e.*DEFT)}\right)^2 = F$$

Step 2: $$\sqrt{\frac{F}{DEFT^2}} = t$$

Adjustment of Means to Control for Background Variation

In recent years, a variety of analyses prepared for the National Center for Education Statistics have employed a strategy of adjusting estimates by covariation among control variables in a given table. Under this strategy, each value of a variable is turned into a dichotomy and regressed on all the other variables under consideration. The parameter values in the regression equation are then used to adjust the estimate. The result basically says, "If the students in the dependent variable evidenced the same configuration of values on the independent variables as everybody else, this is the estimated percentage who would do X."

In the course of this monograph, many studies employing this procedure have been cited (e.g. Cuccaro-Alamin and Choy, 1998; Horn, 1997; Horn, 1998; McCormick, 1997; McCormick, 1999). Given the fundamental question addressed in this study, I have chosen *not* to invoke the "adjustment of means" procedure, rather to let the series of regression equations in Part V tell the story. I do so because in the world subject to practical intervention by the "tool box," people do *not* evidence the same configuration of characteristics or behaviors as everybody else, and the messages one might convey on the basis of adjusted percentages might not be helpful.

For example, in one table, Horn (1997, p. 42) looks at the relationship between high school mathematics course sequence and college enrollment (access) for students with one or more "risk factors" in the NELS-88 cohort. An excerpt from this table should dramatize the case.

High School Math Course Sequence	Unadjusted Percentage of Students Enrolling in Postsecondary Education by 1994	Adjusted Percentage of Students Enrolling in Postsecondary Education by 1994
Algebra I and Geometry	63.1	70.7
At Least One Advanced Course	90.8	73.9

The message to "at risk" students of the "adjusted percentage" is that the highest level of mathematics one reaches in high school really doesn't make that much of a difference in college access. In the real (unadjusted) world, precisely the opposite occurs. Now, which message should we be delivering to all students and particularly to those "at risk"? Enough said.

Appendix B: Inclusions and Exclusions in the Academic Resources Variable: a Demographic Portrait

It is natural to ask whether some population groups are more likely to be excluded from analyses based on the "academic resources" variable, and what one does to rectify the situation. Some analysts, afraid to lose cases from relatively small samples, will impute values on the basis of sub-group means. But I find no persuasive evidence to impute a test score or a high school class rank or the details of a high school curriculum. Instead, one must assume that missing evidence is randomly distributed, and the key to adjustment lies in the various panel weights provided in the dataset. Because exclusions require reweighting, the table below uses unweighted proportions to provide some hints as to how much adjustment for missing evidence a modified weight will account, particularly if the universe is limited to on-time high school graduates.

Who is Missing One or More of the Three Components (Test Score, Class Rank/GPA, Curriculum Intensity & Quality) of "Academic Resources"

Proportion of Students Whose Records Show

	All Three Components	1 or 2 Components	No Components	Event Drop-Outs	H.S. Graduates On-Time	Other
Men	85.4%	13.7%	0.9%	16.2%	82.1%	3.9%
Women	85.5	13.8	0.7	14.7	85.6	4.0
White	88.1	11.5	0.4	14.8	86.3	3.9
Black	82.3	17.1	0.6	18.2	81.6	3.5
Latino	83.1	16.1	0.8	18.3	79.5	4.2
Asian	89.0	11.0	0.0	2.9	90.0	3.1
AmerInd	81.5	17.9	0.6	18.0	74.6	6.8
SES Quintile						
High	88.3	11.5	0.2	5.8	94.8	1.9
2nd	89.5	10.2	0.3	9.1	90.2	2.7
3rd	87.5	12.3	0.2	14.0	87.3	3.3
4th	87.3	12.3	0.4	18.1	82.4	4.6
Low	83.7	15.6	0.7	26.4	73.2	6.7
No HS Diploma	69.3	27.0	3.7	60.3	N.A.	N.A.

Notes: (1) Universe consists of all HS&B/So participants who were alive in 1986. Unweighted N=14,799. (2) "Other" high school graduates include those who received GEDs or other equivalency certificates and those who regular diplomas after 1982. Source: National Center for Education Statistics: High School & Beyond/Sophomores, restricted file. NCES CD #98-135.

It is obvious that the first filter on analysis must be that of high school graduation. Some 60 percent of the permanent ("status") drop-outs (at age 30) were also event drop-outs, and these students tend to be overrepresented in the lowest SES quintiles (not surprising), where black, Latino, and American Indian students are also overrepresented. A relatively high proportion of status drop-outs will not have senior test scores or full high school records. Once one filters the population, it is necessary to create a new senior year weight to account for non-response by reason of missing records.

Appendix C:
Gradations of Academic Intensity & Quality of H.S. Curriculum

The following table sets forth the 40 gradations (in descending value) of curriculum intensity and quality as used in the development of the Academic Resources index and variable. The figures in the boxes represent the minimum rounded number of Carnegie units required for the gradation on a given *row*. Where a box is empty, there are no minimum requirements. Where a box indicates "none" (for remedial math and remedial English), it means that no remedial work is allowed for that gradation. For the High School & Beyond/Sophomore cohort, computer science was not nearly as widely offered as it is today. Therefore, computer related credits were brought into play only to disaggregate lumps in the distribution. Total high school credits/academic credits is an empirically-derived factor that comes into play only in the very lowest gradations.

The basic 5-subject credit thresholds were constructed in the course of examining the coded transcripts for students who were known HS&B/So high school graduates with graduation dates through 12/31/86. The editorial process paid particular attention to all cases that showed 16 or fewer total high school credits. Where the evidence strongly suggested dissonance with other variables in the student's record, all transcripts from that student's *school* were examined. Where non-standard credit metrics were found, they were adjusted, and major components (mathematics, English, etc.) multiplied by as much as (but no more than) 2. The editorial process also Windsorized cases of total Carnegie unit counts above 32, and adjusted the major components down one-by-one.

These gradations of academic intensity and quality are based on the history of one national high school class that was scheduled to graduate in 1982. The next graduating class for which we possess similar national data is that of 1992. While the specific numbers of Carnegie units, APs, and remedial indicators might change, the basic form and principles of the gradations will probably not change. This presentation of the possibilities of high school curricular attainment is criterion-referenced: theoretically, *everybody* can reach gradation level #1.

Grada-tion	English	Math	Science	Langs	Hist or Soc Stu	High Math	Remed Math	Remed English	APs	Comp Science	Total/ Academ Units
1	3.75	3.75	2.0*	2.0	2.0	>A2	None	None	>1		
2	3.75	3,75	2.0*	2.0	2.0	>A2	None	None	1		
3	3.75	3.75	2.0*	2.0	2.0	>A2	None	None	0	>0	
4	3.75	3.75	2.0*	2.0	2.0	>A2	None	None	0		
5	3.0	3.0	2.0	1.0	2.0	>A2	None	None	>1		
6	3.0	3.0	2.0	1.0	2.0	>A2	None	None	1		

(CONTINUED)

Gradation	English	Math	Science*	Langs	Hist or SocSt	High Math	Remed Math	Remed English	APs	Comp Sci	Total/ Academ Units
7	3.0	3.0	2.0	1.0	2.0	>A2	None	None	0	>0	
8	3.0	3.0	2.0	1.0	2.0	>A2	None	None	0		
9	3.75	3.75	2.0*	2.0	2.0	A2	None	None	>0		
10	3.75	3.75	2.0*	2.0	2.0	A2	None	None	0	>0	
11	3.75	3.75	2.0*	2.0	2.0	A2	None	None	0		
12	2.75	2.5	1.0		1.0	>A2			>0		
13	2.75	2.5	1.0		1.0	>A2			0	>0	
14	2.75	2.5	1.0		1.0	>A2			0		
15	3.0	3.0	2.0	1.0	2.0	A2	None	None	>1		
16	3.0	3.0	2.0	1.0	2.0	A2	None	None	1		
17	3.0	3.0	2.0	1.0	2.0	A2	None	None	0	>0	
18	3.0	3.0	2.0	1.0	2.0	A2	None	None	0		
19	3.0	3.0	2.0	1.0	2.0	<A2	None	None	>0		
20	3.75	3.75	2.0*	2.0	2.0	<A2	None	None	0		
21	3.0	3.0	2.0	1.0	2.0	<A2	None	None	0	>0	
22	2.75	2.5	1.0		1.0	A2			>0		
23	2.75	2.5	1.0		1.0	A2			0	>0	
24	2.75	2.5	1.0		1.0	A2			0		
25	3.0	3.0	2.0	1.0	2.0	<A2	None	None	0		
26	2.5	2.0	1.0		0.5	>A2					
27	2.75	2.5	1.0		1.0		Net0			>0	
28	2.75	2.5	1.0		1.0		Net0				
29	2.75	2.5	1.0		1.0		Net1				
30	2.5	2.0	1.0		0.5	A2	Net0			>0	

(CONTINUED)

Gradation	English	Math	Science*	Langs	Hist or SocStu	High Math	Remed Math	Remed English	APs	Comp Science	Total/ Academ Units
31	2.5	2.0	1.0		0.5	A2	Net0				
32	2.5	2.0	1.0		0.5	<A2		None			
33	2.5	2.0	1.0		0.5		Net0			>0	
34	2.5	2.0	1.0		0.5		Net0				
35	2.5	2.0	1.0		0.5		Net1				
36						A2					11/ 6.5+
37									>0		11/ 6.5+
38										>0	11/ 6.5+
39							Net0				11/ 6.5+
40							Net1				11/ 6.5+

Key: A2=Algebra 2; NET1=Total mathematics credits minus remedial mathematics units=0.5 or less; NET0=Total mathematics credits minus remedial mathematics credits=more than 0.5; *=minimum units of Core Laboratory Science, otherwise minimum units of *all* science=2.5.

Appendix D: So They Got a Degree! Why Did It Take So Long? Or Did It?

Both public mythology and public policy are grounded in a belief that a bachelor's degree should be earned within four academic years from the time a student first enrolls in higher education. The four-year benchmark is a standard that assumes continuous enrollment and full credit loads. It derives from a period when relatively few people delayed entry to college to a point in life when they already had families and jobs, and when relatively few transferred from 2-year to 4-year colleges—let alone engaged in any other multi-institutional attendance behavior. Most institutions express the contents of that four-year period in the currency of credits though credits serve far more as proxies for time than content. The result is both a national accounting system and a basis for state funding of public institutions. It is natural for those who plan state budgets to worry about time-to-degree.

The empirical account of time-to-degree is very different from the normative assumption, and only the long-term longitudinal studies of the National Center for Education Statistics can determine that time accurately. No other data bases can do it, and no other evidence is as powerful as 11 or 12 year transcript records, often from two or more institutions in two or more states. The empirical account, measuring *elapsed* time to degree (including periods of stop out, terms of light credit loads, portfolios of credits that were not wholly accepted in transfer—in other words, real life stories), clearly say that the *five-year* bachelor's degree has been the norm since the 1970s. For the generation of the NLS-72 (1972-1984), the mean elapsed time to degree was 54.5 months; for that of the HS&B/So (1982-1993), the mean elapsed time increased to about 57 months. In either case, that's five full academic years.

But there is a different question raised by the multivariate analyses through which we have just passed, and it is a question that should prove very helpful to state, regional, and national planning. Which of the factors that help explain the variance in bachelor's degree *completion* also enlighten our understanding of time-to-degree. Table D1 provides a descriptive exploration of some key variables in pre-collegiate background and attendance patterns in relation both to time and GPA.

Table D3 provides another common and common-sensical account of time-to-degree in the two cohorts of the NLS-72 and the HS&B/So, namely by undergraduate major field and credits earned. This table is slightly different from table D1 because it is not restricted to students for whom a GPA can be computed. In both cohorts, it is obvious that students earning degrees in engineering take longer to do so and earn correspondingly more credits in the course of their undergraduate careers. Since engineering (and architecture, which is included in the same category) programs often require a cooperative semester, these results are not surprising. Education and health sciences/services (nursing, allied health, clinical health science, HPER, etc.) majors in the HS&B/So cohort, however, took much longer to finish degrees and earned significantly more credits than did their counterparts a decade earlier. This change should be subject to further investigation.

Table D1.—Mean elapsed time to bachelor's degree and final undergraduate GPA, by key analytic variables, High School & Beyond/Sophomore cohort, 1982-1993

	Time	S.D.	s.e.	GPA	S.D.	s.e.
ALL	4.72	1.50	.0027	2.89	0.47	.0084
By Number of Institutions						
One	4.26	1.01	.0028	2.92	0.48	.0013
Two	4.87	1.60	.0047	2.88	0.46	.0014
More than two	5.41	1.81	.0071	2.87	0.46	.0018
By Continuity of Enrollment						
Continuous	4.33	0.91	.0017	2.91	0.47	.0087
Non-Continuous	7.12	2.05	.0092	2.82	0.49	.0022
By Academic Resources Quintile						
Highest	4.45	1.24	.0031	3.04	0.46	.0011
2nd	4.81	1.51	.0050	2.80	0.42	.0014
3rd	5.22	1.76	.0089	2.62	0.39	.0019
4th	5.86	2.06	.0176	2.60	0.43	.0037
Lowest	LOW N			LOW N		
By SES Quintile						
Highest	4.60	1.38	.0037	2.92	0.47	.0013
2nd	4.83	1.57	.0055	2.89	0.46	.0016
3rd	4.93	1.67	.0074	2.89	0.49	.0022
4th	4.75	1.55	.0089	2.87	0.47	.0027
Lowest	4.88	1.69	.0154	2.81	0.45	.0041
By Aspirations						
Bachelor's Consistent	4.52	1.31	.0028	2.92	0.46	.0010
Increased to Bachelor's	4.98	1.69	.0056	2.84	0.48	.0016
Lowered from Bachelor's	5.42	1.77	.0109	2.86	0.48	.0030

NOTES: (1) Universe consists of all students who earned a bachelor's degree by 1993, and for whom postsecondary transcripts were received and contained sufficient information to determine elapsed time to degree and undergraduate GPA. Weighted N=806k. (2) Elapsed time is measured in calendar years from the first term of attendance at the "referent" first institution of attendance. (3) Standard errors adjusted by the design effect. Simple s.e.=.675; Taylor series s.e.=1.042; Design Effect=1.54.
SOURCE: National Center for Education Statistics: High School & Beyond/Sophomore cohort, NCES CD#98-135.

At first blush, what we see in table D1 appear to be more confirmations of common sense. First, the number of institutions attended may not be a significant factor in explaining degree completion, but among those who complete degrees, the more institutions, the longer it takes. Second, after the highest quintile of SES, there is no clear relationship between family background and either time-to-degree or final GPA. Third, the variables that contribute most to explaining bachelor's degree completion, academic resources and continuous enrollment, are strongly related to completing degrees in short order. Particularly impressive is the apparent linear relationship between the quintiles of pre-collegiate academic resources and *both* time-to-degree and final GPA.

Table D2.—Factors accounting for completion of a bachelor's degree within a modified traditional time span, High School & Beyond, Sophomore Cohort, 1982-1993

Universe: All students who completed a bachelor's degree, whose transcript records were complete and contain sufficient information to compute time-to-degree and undergraduate GPA. N=3,386. Weighted N=806k. Simple s.e.=.675; Taylor series s.e.=1.042; Design Effect=1.54.

Variable	Parameter Estimate	Adj. s.e.	t	p	Contribution to R^2
Intercept	2.304328	0.2429	6.16		
NOSTOP	0.350468	0.0373	6.10	.001	0.1367
Transfer	-0.177933	0.0380	3.04	.01	0.0517
Freshman GPA	0.117690	0.0287	2.66	.02	0.0230
Low Credits	-0.187478	0.0517	2.36	.05	0.0159
DWI Index	-0.194309	0.0581	2.17	.05	0.0123
Sex	-0.108102	0.0248	2.83	.02	0.0093
Senior Test Quintile	0.026576	0.0152	1.14	---	0.0094
No Return	-0.085257	0.0356	1.56	---	0.0066
Selectivity	0.103977	0.0325	2.08	.10	0.0057
GPA Trend	0.070653	0.0275	1.67	---	0.0040
Number of Schools	-0.060804	0.0287	2.12	.05	0.0027
Remedial Problem	-0.047951	0.0225	1.38	---	0.0025
Doctoral Instit.	-0.053576	0.0273	1.27	---	0.0026
STUWORK	-0.050709	0.0258	1.28	---	0.0022
No Delay	0.103291	0.0530	1.27	---	0.0020
Grant-in-Aid	0.054276	0.0248	1.42	---	0.0014
Credit Ratio	-0.081726	0.0650	0.82	---	0.0011
				R-Sq.	.2890
			Adj.	R-Sq.	.2847

The second blush suggests that a multivariate approach to time-to-degree might tease out the true strengths of competing notions. For this purpose the benchmark became the *median* elapsed time to degree for this cohort, 4.24 calendar years, a figure that translates into about 4.5 academic years. The general form of the question is what background variables, postsecondary attendance patterns, and college performance indicators count most for completing a bachelor's degree in *less than median time*? For purposes of casting a wide net for answers, the academic resources variable (ACRES) was deconstructed into its three components, and all the demographic and attendance pattern variables that were discarded from the models explaining degree completion were brought back. Given the large number of independent variables fed into a single equation, the selection criterion for admission from the correlation matrix was set to the default value of $p \leq .05$. With that setting, table D2 indicates that the following variables have no bearing on time-to-degree: high school class rank/GPA, high school curriculum intensity/quality, race, SES quintile, parenthood, educational aspirations, taking a loan for higher education expenses, and starting at a 4-year college. None of these met the criterion for selection.

Of the 17 variables that met the criterion of selection, only 8 survived in the simple least squares regression model. The only demographic variable in this collection, gender, indicates that males take longer to complete degrees, partly (but not wholly) because men are over-represented in fields where time-to-degree is longer by custom-and-usage, e.g. engineering and architecture.

Table D3.—Time to Bachelor's Degree and Total Undergraduate Credits, by Selected Major, in Two Longitudinal Studies Cohorts

	Time-to-Degree*				Total Undergrad. Credits			
	NLS-72	SD	HS&B/So	SD	NLS-72	SD	HS&B/So	SD
All	4.54	1.53	4.74	1.52	128.9	22.8	135.0	16.9
Business	4.70	1.68	4.73	1.47	126.5	20.3	132.0	14.0
Education	4.43	1.47	5.06	1.75	129.8	21.2	138.8	18.7
Engineering	4.89	1.84	4.93	1.34	139.0	23.8	148.0	19.8
Humanities	4.60	1.52	4.50	1.48	127.3	23.4	129.7	16.0
Arts	4.53	1.33	4.53	1.12	130.0	28.2	136.2	18.1
Social Sci	4.44	1.42	4.60	1.50	124.7	18.1	129.9	14.8
Life Sci	4.46	1.20	4.46	1.08	128.8	22.8	137.5	18.1
Health Sci & Services	4.73	1.53	5.04	1.76	133.8	28.2	143.3	21.1
Physical Sci	4.55	1.53	4.51	1.45	132.5	24.2	137.3	13.1

* In elapsed calendar years.

Appendix E: Example of a Customized National Long-Term Degree Completion Report

There are many ways to cut the national data on bachelor's degree completion. This appendix presents one example of a customized descriptive report. It was originally prepared for the National Center on Public Policy and Higher Education in January, 1999.

For any descriptive report such as this, one must always ask who the "constituent" wants in the denominator of the ratios. In this case, while the data base, the High School & Beyond/ Sophomore cohort, is the same, the resulting universe is different from that used in the text of the "Tool Box" monograph. The degree completion rates, then, are also different—very different. What follows is the report to the Center.

> Students in the universe of analysis were those who (1) had graduated from high school or earned a high school equivalency diploma by the end of 1987; (2) were of known race; (2) attended one or more 2-year and/or 4-year colleges at any time between 1982 and 1993; and (3) earned more than 10 credits. The weighted N is 1,855,866. In the following table, the universe is constricted for selected variables, e.g. SES, highest level of parents' education, etc., where we do not have information for all the students in the sample. The weighted N is stated for these variables.
>
> All differences in columns are statistically significant at $p \leq .05$ unless a pair of estimates is indicated with asterisks. What $p \leq .05$ means is that the odds of the difference occurring by chance are less than one in twenty. This is a standard threshold for significance. This threshold is modified, in the case of each variable, by accounting for the number of categories and comparisons in the variable.
>
> For each variable, the percentage distribution among values is found in the furthest column to the right. For each variable, this column adds to 100.0%. Otherwise, the *rows* add to 100.0, except where affected by rounding.

Proportion at Highest Degree by 1993

	None	Associate	Bachelor	% Of All
ALL	**43.5%**	**11.2%**	**45.2%**	
By Race				
White	39.5	11.3	49.1	81.1
Black	66.6*	8.6	24.8*	10.6
Latino	59.0	15.3	25.8*	5.3
Asian	32.5	9.6+	57.9	2.0
AmerInd	64.2*	14.3+	21.5+	1.0

+These estimates are based on raw numbers that are too low to be reliable.

Proportion at Highest Degree by 1993

By SES Quintile (N=1.686M)	None	Associate	Bachelor	% Of All
Highest	29.7	6.4	63.9	31.3%
2nd	40.8	12.0*	47.2	25.3
3rd	48.5	13.0*	38.4	19.7
4th	53.8	15.1*	31.1	15.0
Lowest	65.4	15.6*	21.0	8.6
By Highest Level of Parents' Education (N=1.752M)				
Bachelor's or Higher	29.7	7.6	62.8	37.9%
Some Postsecondary Ed	46.5	13.7*	39.8	32.2
No Postsecondary Ed	56.2	13.2*	30.6	29.9

[Student reports of the level of parents' education are notoriously inaccurate. In this case, reported level of mother's/father's education was adjusted with reference to reported *occupation*--something late adolescents understand more clearly. Thus, for example, a mother who was reported to be a school teacher but with no education beyond high school had her highest level of education recoded as college graduate. This type of adjustment could be made only for those occupations requiring at least a college education, if not a degree--or a graduate/first professional degree. Since some children have only one parent or report occupation and level of education for only one parent, this variable refers to the highest level of education to be found among any parents.]

By Census Division of Student's High School	None	Associate	Bachelor	
New England	28.1	14.3	57.6	7.6%
Mid-Atlantic	34.7	11.7*	53.6	16.6
East North Central	43.4*	10.9*	45.8*	20.9
West North Central	39.4	13.9*	46.7*	8.6
South Atlantic	44.9*	12.4*	42.8	14.1
East South Central	51.8	10.6*	37.7*	5.2
West South Central	55.8*	7.3	36.9*	9.5
Mountain	55.0*	5.6	39.3*	5.0
Pacific	48.9	11.8*	39.3*	12.6

Proportion at Highest Degree by 1993

	None	Associate	Bachelor	% Of All
By Type of True First Institution Attended				
Doctoral	24.2	2.9*	72.8*	23.8%
Comprehensive	38.3	4.0*	57.7	24.3
Liberal Arts	26.4	3.5*	70.1*	7.0
Specialized 4-Year	40.9	17.4	41.8	2.7
Community College	58.9	21.8	19.3	37.0
Other	70.3	15.2	14.5	5.3

"True First Institution" excludes dual enrollment in high school and false starts.

	None	Associate	Bachelor	% Of All
By Transfer Status (N=1.764M)				
Transferred from 2-year to 4-Year	33.8%	15.0%	51.2%	20.1%
Did Not Transfer from 2-Year to 4-Year	44.5	10.1	45.4	79.9
By Financial Aid, 1982-86				
Grant At Any Time, 1982-1986	36.4	9.7	53.8	47.1%
No Grant, 1982-86	49.9	12.6	37.6	52.9
Loan At Any Time, 1982-1986	34.4	10.3*	55.4	39.9%
No Loan, 1982-86	49.6	11.9*	38.5	60.1

[The years covered by these variables are the four immediately following scheduled high school graduation. Not all students from this cohort who entered higher education were in school during those years.]

Proportion at Highest Degree by 1993

	None	Associate	Bachelor	% Of All
By Work-Study, 1982-86+				
Every Year, 1982-86	11.6	Low N	83.5	6.2%
Some Years, 1982-86	38.1	12.6*	49.4	36.7
Never, 1982-86	23.3	Low N	72.0	4.4
Did Not Work to Help Pay for Education or Worked Only During Summers	52.8	11.6*	35.6	52.7

[+A difficult category because the loop of questions was confined to students who said they worked to contribute to their education. "Work-Study," as defined here, excludes summer jobs and covers only that work taking place while the student was in school.]

ADDENDUM: States in Census Divisions*

New England: ME, NH, VT, MA, RI, CT

Mid-Atlantic: NY, NJ, PA

E. North Central: OH, MI, IN, IL, WI

W. North Central: MN, ND, SD, IA, MO, KS, NB

South Atlantic: DE, MD, DC, VA, WVA, NC, SC, GA, FL

E. South Central: KY, TN, AL, MS

W. South Central: LA, AR, TX, OK

Mountain: MT, ID, CO, NM, UT, AZ, NV, WY

Pacific: CA, OR, WA, HA, AK

*Divisions used by the Bureau of the Census in the Current Population Series (CPS).

☆ U. S. GOVERNMENT PRINTING OFFICE: 1999-452-108/10323

United States
Department of Education
Washington, DC 20208–5531

Official Business
Penalty for Private Use, $300

Postage and Fees Paid
U.S. Department of Education
Permit No. G–17

Standard Mail (B)

PLLI 1999–8021

NOTICE

REPRODUCTION BASIS

☐ This document is covered by a signed "Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☑ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").