federal register

THURSDAY, JUNE 17, 1976

PART II:

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of the Secretary

■

PRIVACY RIGHTS OF PARENTS AND STUDENTS

Final Rule on Education Records

(
tion
infc
of a
of tl
ble s
(1.
ing t
stitut
who l
or ins
tional
(2)
school
or inte
quirem
(3) S
in § 99..
of:
(i) Tl
United S
(ii) T.
(iii) T
of the N
or the As:
or
(iv) St:
(4) In
for which :
a student
personally
the educati
be disclosec
for such pu
(i) To det
student for 1
(ii) To de
financial aid.
(iii) To (
which will be
nancial aid, o
(iv) To en
tions of the fir.
(5) To Stat
thorities to wh
cally required 1
pursuant to St.
to November .
graph applies o
quire that spec
closed to State c
not apply to sta
do not require
this paragraph
from further limi
of State or local
closures are mad
graph;
(6) To organ
studies for, or on
agencies or institu
of developing, val
tering predictive
student aid progra
instruction; *Provide*
are conducted in a
not permit the perso
students and their pa
other than represent
nization and the info
stroyed when no lon
purposes for which tl
ducted; the term "c
cludes, but is not lir.
State and local agenci
ent organizations

RECORD OF DISCLOSURE

DATE:

SCHOOL NAME:

STUDENT NAME:

The Commissioner of Education has directed the Research Triangle Institute to conduct an evaluation of the Upward Bound program and has requested a copy of this student's transcript for use in this evaluation.

This request is in conformity with applicable State and Federal privacy laws and regulations. For further information, please contact the U. S. Office of Education (202-245-8057) or the Research Triangle Institute (919-541-6000).

Graham J. Burkheimer
Project Director
Research Triangle Institute
Post Office Box 12194
Research Triangle Park, N. C. 27709

RALEIGH INSTITUTE OF TECHNOLOGY
2505 15TH STREET
RALEIGH, NC 27608

THE STUDENTS LISTED BELOW HAVE INDICATED THAT THEY ATTENDED YOUR INSTITUTION ON OR AFTER THE DATES INDICATED. PLEASE PROVIDE TRANSCRIPTS FOR STUDENTS LISTED. ALL STUDENTS LISTED, OR THEIR PARENTS, HAVE BEEN NOTIFIED OF OUR INTENTION TO OBTAIN TRANSCRIPTS AND GIVEN THE OPPORTUNITY TO REFUSE ACCESS. STUDENTS' WISHES IN THIS MATTER ARE BEING HONORED. A VOUCHER FOR REIMBURSEMENT, RETURN ENVELOPE, AND RECORDS OF DISCLOSURE FOR ALL LISTED STUDENTS HAVE BEEN PROVIDED.

STUDENT NAME	APPROXIMATE DATE OF LAST ENROLLMENT
JOHN SMITH	FALL 1978
JANE JONES	SPRING 1977
PAUL GREEN	FALL 1977

RESEARCH TRIANGLE INSTITUTE
POST OFFICE BOX 12194
RESEARCH TRIANGLE PARK, NORTH CAROLINA 27709

RTI

CENTER FOR EDUCATIONAL RESEARCH AND EVALUATION

May 10, 1979

Dr. John Jones
Any University
Any Street
Any Town, Any State 00000

Dear Dr. Jones:

Enclosed you will find the package of materials that we discussed with you during our recent telephone conversation. As explained in that conversation, we originally mailed a similar package to your school in March. Since you did not receive it, we are sending a replacement set with this letter. The enclosed materials include: a letter from the Deputy Commissioner for Higher and Continuing Education, pertinent sections regarding privacy regulations as related to Education Records, a legal interpretation from the Office of General Counsel, a statement of privacy regulation procedures implemented for this study, a listing of students for whom transcripts are requested, a record of disclosure for the file of each student listed, and a voucher to facilitate payment to your institution for the cost incurred in providing the transcripts requested. Again, we regret any additional inconvenience caused your office due to your not receiving our original request.

Thank you for your cooperation and timely attention to this matter. Should you have any additional questions, please feel free to call me or Dr. John Riccobono (toll free) at 1-800-334-8571.

Very truly yours,

G. J. Burkheimer, Jr., Ph.D.
Project Director

GJB:bge

Enclosure

Appendix E

Data Collection and Processing

Appendix E

Data Collection and Processing

I. GENERAL

This appendix describes the data collection and processing procedures for the second follow-up study of the UB program. Instrument development, mail and telephone survey activities, acquisition and use of extant data files, receipt control, and manual and machine processing procedures are described in detail.

The remainder of this introductory section provides a general description of data collection and processing, an overview of quality control procedures, and a description of procedures used to ensure data confidentiality. Section II provides a description of the several data sources used. Section III details all survey data collection activities and provides response rate information. Section IV describes the survey control system that was used in this study. Section V considers manual processing activities for questionnaires and transcripts. Section VI presents the details of machine processing of data, and Section VII provides information relating to the quality of student data.

A. Overview of Data Collection and Processing

The second follow-up study of the UB program was a continuation of a data collection effort for the sample of UB and comparison students who were selected for participation in the base-year survey (see Appendix C, for base-year and follow-up sampling methodologies). In preparation for second follow-up data collection, three distinct but related efforts to update directory file information (i.e., name and address) for all eligible sample members were conducted. The first effort was initiated through BHCE/USOE and entailed a request to UB projects involved in the base-year study to initiate local tracing operations toward obtaining available directory file updates. The second approach was initiated at RTI, and involved mailing a newsletter (containing relevant findings from the base-year and first follow-up studies) and a request for corrected directory file information to all study sample members (see Appendix D). Newsletters that were returned by the postal service as undeliverable generated telephone tracing activities in an attempt to locate the sample members. A final effort for those not located through previous steps was a request for directory information from Departments of Motor Vehicles of the state of last known address.

Data collection in the second follow-up survey was accomplished in two stages. The first stage involved a mail survey, using a brief self-administered questionnaire, of all eligible sample members. This mail survey consisted of an initial questionnaire mailout, followed one week later by a mailed thank-you/reminder postcard. About four weeks after the initial questionnaire mailing, a second questionnaire was sent to all who had not responded by that time. Four weeks following the second mailing, a probability sample of 1,500 mail survey nonrespondents was selected for subsequent telephone interview to gather the information requested on the questionnaire. The second stage of data collection involved obtaining academic transcripts from the PSE institutions currently or last attended by sample members, as indicated by the completed questionnaires or telephone interviews. An Institutional

Request Package was developed and mailed to all identified postsecondary institutions.

In addition to data obtained from the second follow-up questionnaires and student PSE transcripts, other data were obtained from existing sources. Student background information, prior education and work status, and prior UB participation status were available from a student master file comprised of data collected or constructed during the base-year and first follow-up studies. Information characterizing the PSE institutions attended by sample members was obtained from an institutional data file provided by USOE.

The operational phase of the second follow-up study covered the period from April 1978 through September 1979. Major activities during this period involved data collection, processing and analysis. To facilitate data collection and processing operations, the entire survey was monitored by a computerized control system from the time of first newsletter mailing until an edited machine-readable data base was established. Mailed questionnaire returns were manually edited for completeness and consistency, and returned instruments that failed the initial edit triggered telephone calls to the sample members to clarify or complete the questionnaire responses. Data were then coded into machine-readable form by programmable key-to-tape equipment. Similar procedures were employed for transcript data.

Additional processing of data was accomplished by computer. Data were thoroughly edited (including checks for out-of-range responses, checks for inconsistent responses, and implementation of imputation procedures). Data were subsequently merged into a master data file (which included data from the base-year and first follow-up study). Similar processing procedures were employed for transcript data. Finally, data from the institutional characteristics file were merged with those collected through the second follow-up questionnaire and institutional request surveys and analytic variables and weights were developed.

B. Overview of Quality Control Procedures

RTI was concerned throughout the period of data base construction with the quality of data entered. Quality control of data collected during the second follow-up study occurred at four levels. The first level was an overall data management process implemented through the computerized survey control system. Use of this system considerably reduced errors related to inappropriately identified data (faulty ID numbers). The second level of quality control was a scan-edit of incoming data for completeness and consistency of critical data items, followed by telephone contacts to resolve incomplete or inconsistent responses. The third level of quality control existed at the stage of keying data. Data were keyed through key-to-tape consoles programmed for routine data checks, including: ID number check digit verification, range checks, and consistency checks. Moreover, a 10 percent sample of all data were rekeyed by two additional operators as a check for possible keystroke error. The fourth and final quality control level occurred at the time of entry of data into the computer data file. Machine edit of data involved more sophisticated checks for range and internal consistency as well as consistency checks between data sets (i.e., consistency of questionnaire responses and transcript data, and comparison of information from both of these sources with data obtained previously during the base-year and first follow-up studies).

C. Confidentiality of Project Data

All data collected in the study were maintained under strict control to insure confidentiality of individual data and to maintain data integrity. RTI's Committee for the Protection of Human Subjects reviewed second follow-up planning documents and maintained involvement in all relevant aspects of the survey for the protection of the Institute, USOE, and the study participants.

RTI assumed the responsibility for maintaining all project documents--including questionnaires, transcripts, and data tapes--under strict control to ensure not only confidentiality but also the integrity of various data. During all stages of processing and storage, senior project personnel controlled access to and removal and replacement of all documents from specified working and storage areas.

Second follow-up materials were received in the mail on a daily flow basis and, insofar as possible, were dealt with on a flow basis. Overnight storage of processing batches was in a locked and secure work space. When processing was completed, source doucments were filed in an ordered, accessible manner in a controlled-access storage space.

The name and address file that was necessary for distribution purposes and the substantive data files required for analysis purposes were maintained separately on a strict need-to-know basis throughout the duration of the project. At no time were these files readily accessible except under specific authority of a senior professional working on the project.

II. DATA SOURCES

The data used in the second follow-up study came from six logically separate data sets: (1) an extant Student Master file comprised of data collected or constructed during the base-year and first follow-up studies; (2) the second follow-up survey control file; (3) responses to the second follow-up questionnaire; (4) an institutional characteristics file of 11,010 PSE institutions, developed and provided by USOE;[1] (5) listings of funded Trio programs from 1974 through 1979;[2] and (6) data abstracted from student transcripts, obtained from the PSE institution presently attended or last attended by sample members. Because extant data sources were used, additional instrumentation for the second follow-up study was restricted to the student survey instrument; however, each data source is discussed below.[3]

---

[1] See C. D. Carroll, Characteristics of Postsecondary Education Institutions: Computer File Version 7907. USOE Technical Paper 79-12, Washington, D.C.: U.S. Office of Education, June 1979.

[2] See, for example, Johnson, D. D. and Sedorchuk, F. Directory of Funded Projects: Program Year 1978-79 (FY '78 Funds). Washington, D.C.: U.S. Office of Education, October 1978.

[3] See also Burkheimer, G. J., Richardson, J. E., and French, A. M. Data Base Design for the Second Follow-Up Evaluation of the Upward Bound Program. (Report No. RTI/1605/06-04S) Research Triangle Park, N.C.: Center for Educational Research and Evaluation, Research Triangle Institute, October 1979.

A. The Second Follow-up Questionnaire (SFUQ)

The SFUQ was specified and designed by RTI in close collaboration with OED/USOE staff members. In general, the SFUQ was intended to collect data relevant to the student outcome constructs specified in the analysis questions presented in Chapter 1 of this report, with a minimum burden on the respondent in completing and returning the questionnaire. RTI conducted clinical pretests of initial versions of the survey questionnaire, using less than ten students from local PSE institutions. The purpose of these pretests was to investigate various item formats, as well as overall instrument format, for clarity and ease of completion. The questionnaire was revised as necessary, based on reviews of the pretest results by RTI and other interested persons at USOE.

Two forms (A and B) of the SFUQ were developed and designed for self-administration by the respondent. Form A contained 15 items and Form B contained 16 items. The two questionnaire forms were identical in terms of format and content with one exception: Form B contained an additional question concerning extent and pattern of UB participation to supplement the data base where data on this variable were not obtained previously. Thus, Form A was intended for sample members for whom information regarding prior UB participation was available, whereas Form B was intended for those for whom such information was missing.[4]

SFUQ items (on both forms) dealt with factors concerning the individual's current education and work status; educational history (types of schools attended and approximate dates of attendance); educational aspirations and expectations; sources, amounts, and perceived adequacy of financial support; and perceived deterrents to educational attainment. Almost all of the questions on the SFUQ were of the forced-choice (closed-response) type. Open-ended, or free-response, items were limited to questions involving name and location of postsecondary school currently or last attended and to two question s related to financial aid.

Cover letters explaining the purpose of the study and soliciting cooperation, and an instruction sheet providing directions for completing the questionnaires, were also prepared to accompany the instrument on each of several possible mailings. Copies of both forms of the SFUQ, with cover letters and instructions, are included in Appendix D.

B. The Institutional Request Package (IRP)

A major purpose of the second follow-up study was to investigate the impact of the UB program in terms of the PSE performance of former participants. To accomplish this objective, student PSE performance data (i.e., information pertaining to the individual's major field of study, grade point average, credit hours attempted/completed, etc.), as recorded on academic transcripts, were collected from the PSE institutions currently or last attended by sample members. Toward this end, the IRP was developed for mailing to the PSE school currently or last attended by each sample member (as indicated by response on the Second Follow-up Questionnaire).

---

[4] By error, Form A was mailed to former participants only and Form B was sent to nonparticipants. When this error was discovered, the UB participation data for former participants who should have received Form B was collected by telephone.

The IRP consisted of a cover letter and several enclosures. The cover letter briefly described the purpose of the study, stressed its importance, informed the administrator that RTI was conducting the survey for USOE, informed the institution of the contractor's rights of access to the information by referencing current Federal privacy legislation (enclosed separately), assured that such legislation would be scrupulously observed, and requested prompt cooperation of the school in this study. The letter further explained what was required of the school and informed the administrator that the standard transcript delivery fee would be paid to the school to offset reproduction and related clerical expenses.

Also included in the IRP were: a computer-generated form containing a complete list of students who indicated attendance at that school and for whom transcripts were requested, with their approximate date of last enrollment; a statement of procedures employed to insure protection of the individual's data; pertinent sections of the current Federal Privacy Regulations; an interpretation of the Regulations, as applicable to this study, by the Office of the General Counsel, HEW; a voucher for reimbursing the school for incurred expenses; a "record of disclosure" for placement in the file of each involved student stating that RTI/USOE had accessed the student's file; and a postage-paid return envelope. A copy of the IRP is included in Appendix D.

C. Other Data

Additional data for the second follow-up were available from the base-year and first follow-up surveys. These data had been incorporated into a student master file, which has been documented separately.[5] The information obtained from this file consisted of encrypted student identification numbers, base-year, and first follow-up sampling parameters (including raw sampling weights), individual response histories for the base-year and first follow-up studies, selected student characteristics and analytic variables determined during the base-year study, and analytic variables and analysis weights developed during the first follow-up study. Major analytic variables extracted from the file were: (a) student classifying information (age, initial grade cohort, extent of prior UB participation, sex, race, academic risk status, poverty status), and (b) prior educational status in fall 1973, spring 1974, fall 1974, and spring 1977.

An extant Institutional Characteristics File was created by USOE. Original plans had called for exclusive use of an earlier version of the institutional file; however, when data became available over 300 identified institutions were not represented on this earlier version. At that time, USOE provided the newly-developed file, which contained sufficient information to compute most of the previously defined institutional indices to be used in this study. Data to complete the Institutional Indices for all schools were obtained, to the extent possible, from extant data of record, and added to the file for identified schools. Some data were imputed, for those cases in which the school name was indicative of certain school characteristics. For important

---

[5] Research Triangle Institute, Upward Bound Follow-Up Evaluation Analysis File: Standard Codes and Tape Layout (Project 22U-889). Research Triangle Park, N.C.: Center for Educational Research and Evaluation, Research Triangle Institute, September 1977.

institutional indices, involving the UB host status of institution and Special Services for Disadvantaged Students (SSDS) host status of the institution, data were not provided on the institutional characteristics file; however, these indices were obtained from extant data of record (program directories) and added to the data file. These data had already been encoded for most of the involved schools on the earlier version of an institutional file.

III. SURVEY OPERATIONS

This section briefly describes the various operations employed in collecting and processing data for the second follow-up study, as well as the results of those operations. Continued participation of the original sample members in this evaluation study was largely dependent upon RTI's success in locating, motivating, and establishing rapport with these individuals. Materials used by RTI to enhance cooperation and increase response in the second follow-up included newsletters, $2.00 incentive checks, thank you/reminder postcards, and prompting letters. Additional methods involved intensive efforts at tracing sample members to obtain current directory information. The details and results of these data collection activities are presented in this section. Survey operations were monitored by a computerized survey control system, as described in Section IV below.

A. Preliminary Data Collection Activities

Although actual data collection for the second follow-up began in October 1978, preparation for this work began several months earlier. The first major activity involved updating the directory file for all eligible sample members. This file was maintained in a password-protected computer file and contained last known name, address, telephone number, and encrypted ID number for each member of the base-year study sample. An additional access-controlled paper file was maintained for all sample members containing results of previous tracing activities and old addresses.

The base-year study involved a sample of 6,050 UB and comparison students; however, 114 of these original sample members had been classified as ineligible after completion of the first follow-up study due to their inappropriate definition in the initial sampling frame, death, or prior refusal. These ineligibles were excluded from the tracing activities and from the second follow-up survey. Some initial updates to the directory file were made on the basis of address corrections provided by students so late in the first follow-up study that the information had not been previously processed. Major additional activities relating to updating the directory file, involving BHCE/USOE-tracing, newsletter mailings, telephone tracing operations, and additional tracing through Departments of Motor Vehicles, are shown in Figure E.1 and are discussed below.

1. BHCE/USOE Tracing

The initial directory file update effort was conducted by BHCE/USOE. At BHCE/USOE request, RTI generated lists, by project, of last known names, addresses, and telephone numbers of individuals for whom no response to the first follow-up was received or for whom the first follow-up questionnaire was returned by the postal service as undeliverable. BHCE/USOE then forwarded these lists to the directors of the local UB projects with which the sample

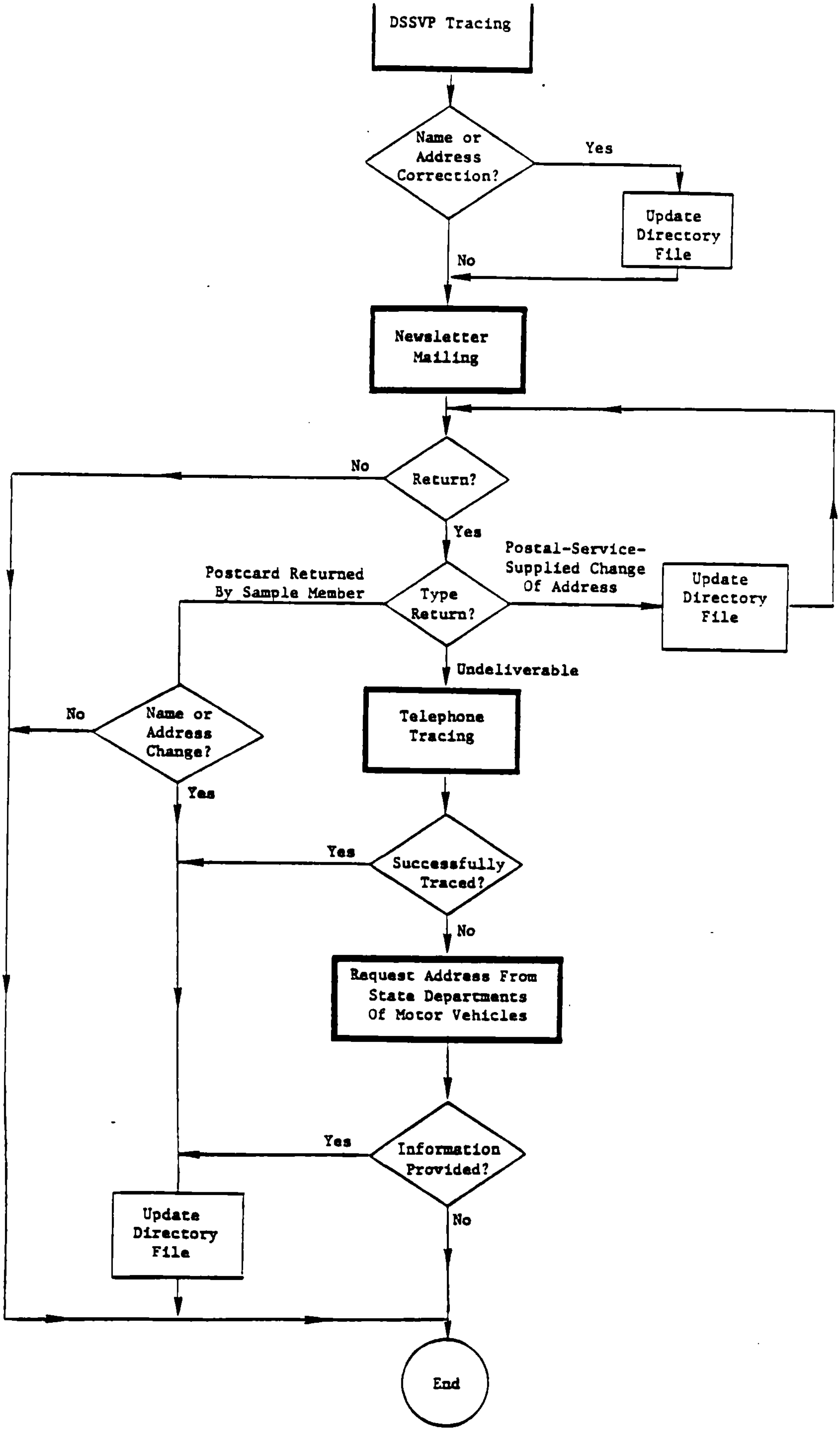


Figure E.1. General flow of tracing activities for the Upward Bound Second Follow-up study.

members had been associated.[6] Project directors were requested by BHCE/USOE to attempt to obtain correct directory information for all listed individuals. Updates received from these projects were entered into the directory file on an as-received basis.

2. Newsletters

As a second directory file update effort, an initial second follow-up activity involved the preparation of two distinct newsletters that were mailed (forwarding address correction requested) to all eligible former UB participants and nonparticipants in the original sample. The two newsletters, one for former UB participants and one for nonparticipants, presented selected results of previous surveys, announced that the second follow-up survey was underway, stressed the importance of continued participation, and requested that the individual return an enclosed postcard verifying or updating his or her name, address, and telephone number.

Experience with the use of an advance newsletter in other large-scale national surveys conducted by RTI had shown this procedure to be extremely useful for initiating tracing activities well ahead of questionnaire mailout as well as for establishing continued rapport with sample members. Copies of both newsletters with accompanying cover letters and return postcard are included in Appendix D. Mailing of newsletters to the last known address of all eligible sample members was begun on 28 April 1978 and continued through 3 July. The extended mailing period included delayed mailings and/or remailings to 563 individuals for whom correct mailing addresses were obtained (through various tracing activities) subsequent to the first mailing. As a result of the various newsletter mailings, 1,774 sample members returned the newsletter postcard either verifying or indicating changes in name and/or address. In addition, the postal service returned 679 newsletters as undeliverable, and these were sent to RTI's telephone tracing department for follow-up action.

3. Telephone Tracing

Telephone tracing procedures were implemented if no current address was available for an individual or as soon as a newsletter was returned by the post office indicating that the mailing address for the individual was no longer correct and that no forwarding address was available. Tracing information types and sources available from the previous surveys were: (a) last known address of respondent; (b) name, address, and telephone number of parents, guardians, or relatives; (c) name and location of the high school or postsecondary school the individual attended; (d) UB project (if applicable); and (e) neighbors or friends of the individual. Receipt of initial directory information changes and tracing of undeliverables continued through September 1978. As new addresses were obtained from tracing activities, returned mail, or other sources, the computer directory file of names and addresses was updated. If this information had been obtained prior to 3 July, a newsletter was mailed to the individual.

---

6 Nonparticipants were also associated uniquely with one and only one UB project by virtue of being associated with a high school served by the project.

4. Additional Tracing

After previous tracing activities, there still remained 357 cases for whom no current address or only an incorrect (undeliverable) address was available. It was jointly determined by RTI and OED that one additional attempt would be made to obtain current addresses for these individuals. That attempt was made through the Department of Motor Vehicles in the state of last known residence for the 170 individuals for whom necessary information to initiate such tracing (e.g., birth date and/or social security number) was available. This additional tracing was initiated on 3 August 1979, and resulted in updated addresses for 56 individuals.

5. Overall Tracing Results

As a result of the various tracing activities, more current name and address information was obtained for 1,622 individuals. (An additional 935 individuals verified that the current name and address information maintained in the RTI directory file was correct.) Of these cases, 839 represented return of the postcard included with the newsletter, 613 represented cases resolved through one of the tracing activities (DSSVP, telephone, and Departments of Motor Vehicles), and 170 represented only a postal-service-supplied change of address. Further, as a result of the various tracing activities 9 sample members were identified as deceased or otherwise incapacitated.

B. Student Survey

The student survey consisted of several stages, including an initial mailing of the SFUQ, a second mailing to nonrespondents to the first mailing, remailings to traceable cases for whom the first or second mailing was returned as undeliverable, and telephone interviews of a subsample of nonrespondents and undeliverables. The general nature of this data collection effort is shown in Figure E.2 and discussed more thoroughly in the following subsections.

1. Mail Questionnaire Survey

As noted earlier, 114 of the original 6,050 student sample members selected for participation in the base-year study had been classified as ineligible for further study after completion of the first follow-up survey. Additionally, the 9 sample members determined to be deceased during tracing were classified as ineligible. Another group of 53 individuals was not included in the mailing, since their address had never been determined. Thus, only 5,874 sample members were involved in the survey. The initial packet of survey materials was mailed on 10 October 1978, and included a SFUQ with instructions, a cover letter from the Deputy Commissioner of the Bureau of Higher and Continuing Education, and a $2.00 incentive check. One week later, thank-you/reminder postcards were mailed to these same individuals as a prompting for questionnaire return. Mail returns consisted of (a) packages returned by the postal service as undeliverable, (b) postal-service-supplied forwarding addresses, and (c) completed or partially completed questionnaires. As returns arrived, they were event coded for entry into the survey control system, name and/or address corrections were made in the directory file, and telephone tracing activities were initiated for undeliverable cases.

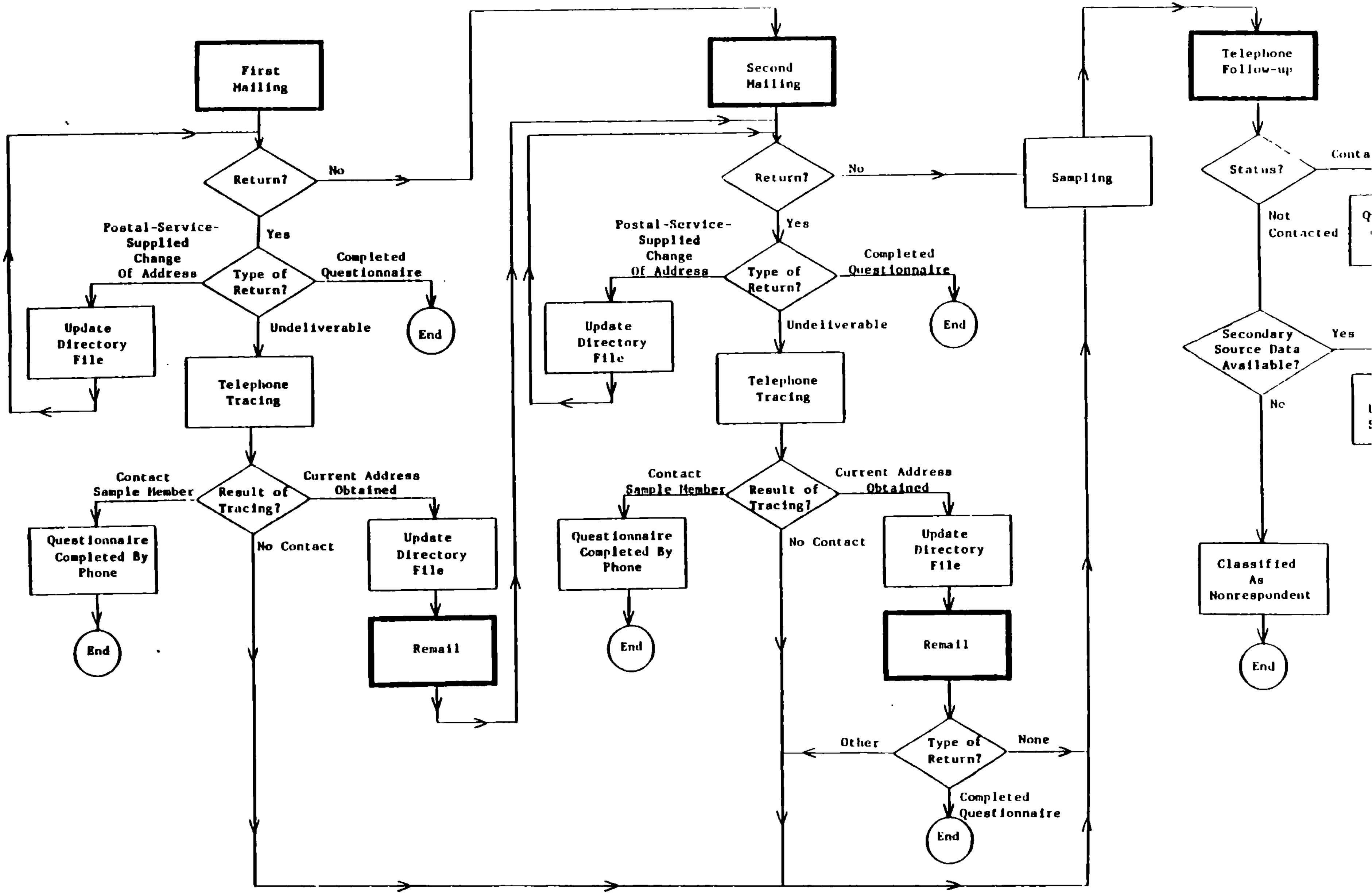


Figure E.2. General flow of student questionnaire data collection activities for the Upward Bound Second Follow-up study.

On 6 November a second questionnaire packet--containing the same materials as the first, except for the $2.00 incentive check, plus a prompting letter from the RTI project director--was mailed to 3,744 nonrespondents to the first mailing. "Remailings" were conducted on 10 November and 6 December for sample members whose first or second questionnaire packet had been returned as undeliverable but who had been subsequently traced. These remailings consisted of the same materials included in the first mailing, along with a special explanatory letter from the RTI project director. Each of these mailings was followed one week later by a thank-you/reminder postcards to those individuals. Materials included in the several mailings are included in Appendix D.

Returns to the various mailings and remailings continued to arrive at RTI through February 1979. A total of 3,005 questionnaires were received from the 5,874 eligible sample members involved in the mail survey (excluding those who returned questionnaires and were also selected into the telephone interview sample), for a 51.2 percent mail return rate.

2. Telephone Interviews

The second follow-up study design incorporated a telephone follow-up survey of those individuals who had not responded to the mail questionnaire by 3 January 1979. It had been anticipated that mail respondents would differ from nonrespondents in terms of the outcome measures being assessed by the study, and coverage of this segment of the original sample was considered quite important. Since cost considerations precluded contacting all mail nonrespondents, a stratified subsample of 1,500 mail nonrespondents was selected for telephone follow-up.

Based on suspected biases, expected standard errors, analytic appropriateness, and expected costs, an optimal (or near optimal) sampling procedure was determined jointly by RTI and USOE. Individuals who had been classified as "hard core" undeliverables (N = 308) were not included in the sampling frame; they were considered lost to the study, and their weights were appropriately redistributed (see Appendix G). Ten strata were defined by the sampling frame, based on a partial crossing of: types of nonresponse (undeliverable or simply unreturned questionnaires), availability of first follow-up data (available or not available), and prior indication of PSE entry (previously indicated entry, previously indicated no entry, no previous indication). Differential sampling rates were established within each strata with higher rates being assigned primarily to prior PSE entry (in keeping with the major objectives of the study) and to simple nonrespondents (due to heavy cost factors estimated for contacting undeliverables). Since no serious response rate differences were indicated among other categories of interest (i.e., former participation status, initial grade cohort, UB project membership, race, sex, poverty status, or risk status), these categories were not used to define formal strata; however, the sampling procedure used (systematic selection) did take these factors into account to assure stochastic representation among the various categories. The sampling procedure thus ensured that all eligible sample members (excluding the "hard core" undeliverables) had a positive probability of representation in the final data set. A more detailed statement of the sampling design employed for selecting nonrespondents to the mail survey is presented in Appendix C.

Telephone tracing and interviewing of those selected occurred from 11 January through 16 February, and involved attempts to locate the previous nonrespondents and to collect, by telephone, the basic information on the SFUQ. During the tracing process, questionnaire information about the sample member was solicited from family members and other secondary sources. Any information so collected was used if the sample member could not be personally contacted; however, obtaining these secondary source data in no way restricted further attempts to contact the sample member.

To eliminate unnecessary expense and unnecessary contact with sample members or their families, mail returns were checked against current lists of telephone interview completions; and, if a mail return was received for a sampled individual prior to completing telephone interview procedures, subsequent telephone activities for that person were suspended. Some mail returns were not received in time to avoid duplication of data collection efforts, but a substantial number were received prior to telephone interview. Of the 1,500 individuals selected for telephone interview, information was collected (from primary or secondary sources) on 1,259, for an overall telephone completion rate of about 84 percent.

3. Overall Survey Return Rates

Information was obtained for 4,264 of the 5,874 individuals included in the second follow-up survey (3,005 mail returns and 1,259 telephone interviews), for a total survey response rate of about 73 percent. Sample coverage was considerably higher (88 percent) due to the high contact rate among the mail nonrespondents subsampled for telephone interviews. Coverage rate is still low, however, since it reflects the fact that "hard core" undeliverables (N = 308) were not included in the telephone interview sampling frame. Final return rates for the various data collection efforts (mail and telephone surveys) are given in Table E.1.

Final data availability in relation to the major study classifiers (UB participation and initial grade-level cohort) is given in Table E.2. The values presented in Table E.2 indicate some differential data availability. In general, data availability was slightly greater for the former UB participant group than for nonparticipants. Differences in data availability among the three grade cohorts were larger and were systematically increasing over the three grades considered.

Sample coverage was sufficiently complete to assure that any biases in the findings, due to peculiarities of the nonresponding group, would be relatively small. The greater availability of data for former participants, however, raises the possibility that those individuals representing former participant nonrespondents are in some way systematically different from those representing the nonparticipant nonrespondents. Cohort nonresponse differences also may have led to some distortion of results, although this variable was typically controlled in analysis. Similarly, the lower completion rate of telephone interviews among sample members for whom current addresses were not available may also have distorted the results. Statistical adjustment techniques were applied in an attempt to correct for any biases that may have been introduced by differential data availability (see Appendixes G and H).

Table E.1

FINAL RESPONSE RATES FOR THE UPWARD BOUND SECOND FOLLOW-UP STUDY

Group	Number	Percent of Eligibles
Ineligibles	123	--
Nonrespondents[a/]	1,660	28.0%
Respondents		
Mail 1 (M1) only	2,388	40.3%
Mail 2 (M2) only	532	9.0%
Telephone (T) only	1,243	21.0%
M1 and M2	86	1.5%
M1 and T	10	0.2%
M2 and T	7	0.1%
M1, M2, and T	1	0.0%
Total Respondents	4,267	72.0%
Total Eligibles	5,927	100.0%
Total	6,050	

NOTE: A total of 11 individuals represented as mail respondents represent cases, for which data was collected by telephone during tracing. Three deceased sample members are reported as respondents. Telephone response rates include 393 secondary source responses.

[a/] Includes all "undeliverables," including "hard core" cases.

C. Institutional Requests Data Collection

As noted earlier, second follow-up data collection also involved a separate request for transcripts from PSE institutions most recently attended by those sample members who reported PSE attendance. Of the 4,264 eligible sample members responding (mail and telephone) to the Second Follow-up Questionnaire, 3,090 indicated current or prior PSE attendance at 999 different institutions. Of these, academic transcripts were requested for only 2,656 individuals at 873 PSE institutions. Transcripts were not requested for the remaining 434 sample members for the following reasons: (1) the respondent explicitly refused to grant permission to access his or her transcript data (N = 259); (2) the PSE institution last attended was not identified (N = 68); or, (3) no current mailing address could be established for the PSE institution identified (N = 107).

Table E.2

DATA AVAILABILITY WITH RESPECT TO PRIOR UPWARD BOUND PARTICIPATION STATUS AND ORIGINAL GRADE COHORT

Fall 1973 Grade Level[a]	Upward Bound Participant Status[b]			
	Prior Participant		Nonparticipant	
	Number of Cases[c]	Data Available[d]	Number of Cases[c]	Data Available[d]
10	552	366 (66.3%)	764	515 (67.4%)
11	1,393	1,037 (74.4%)	770	518 (67.3%)
12	1,675	1,265 (75.5%)	770	563 (73.1%)
TOTAL	3,620	2,668 (73.7%)	2,304	1,596 (69.3%)

a/ Classification used is best estimate of original grade cohort.

b/ Classification used is that used for original sampling.

c/ The 126 cases who were ineligible for the survey or who were reported to be deceased are not included in these counts.

d/ Percentages expressed are relative to number of cases in the specific cell.

The survey plan had called for mailing of institutional request packages to all PSE institutions simultaneously. However, due to problems in identifying a number of the institutions involved at the scheduled mailing time, the transcript request survey was conducted in two waves. More specifically, 331 PSE institutions (representing 433 students) were unidentifiable through any of the available directories at the scheduled mailout date. To avoid jeopardizing the overall project schedule, it was decided that further efforts would be made to identify as many of the unidentified schools as possible and to make transcript requests of those identified subsequently. These efforts were successful in appropriately identifying and/or obtaining directory information for 239 of the previously unidentified schools (representing 326 students). On the other hand, addresses could not be obtained for 92 PSE institutions and, therefore, it was not possible to initiate transcript requests for the 107 sample members who had reported current or last attendance at these schools.

On 16 March, the first-wave mailing of transcript request packages was sent to 634 postsecondary institutions. Two weeks later, on 19 March, mailings were made to the 239 institutions in the second wave. Copies of materials included in the transcript request package (first and second waves) are included in Appendix D.

Responses to the transcript requests continued to arrive at RTI through May. These responses were routinely logged (by both student and institution), event coded, batched, and entered into the survey control system. Vouchers for payment to institutions for transcripts were processed as received to facilitate timely payment. Receipt of mail and telephone inquiries from schools were handled individually on an as-received basis. Most of these inquiries were either to clarify transcript release policy or to seek additional identifying information (e.g., social security number) on students for whom transcripts had been requested.

Telephone follow-up activities were intiated on 16 April to 194 institutions that had not responded to the first-wave transcript request mailing, and on 3 May to 72 nonresponding institutions in the second-wave mailing. The purpose of these contacts was: (1) to determine if the institution had recieved the request, (2) if so, to urge early response, and (3) if not, to obtain the name of an individual to whom a second package could be mailed. Remailing of institutional request packages to specified individuals in the latter schools was initiated on 30 April and 16 May.

Responses (either partial or complete) were received from 817 (over 93 percent) of the institutions requested and transcripts were received for 2,310 sample members (87 percent of those requested). An additional 190 sample members at 127 of those institutions (about 7 percent) were reported as never having been enrolled in the school.

IV. THE SURVEY SUPPORT SYSTEM

Survey support includes a broad range of management activities directed toward providing control and aid to other survey operations. The basic survey support tasks were: (a) directory file maintenance, (b) sample selection, (c) mailing label production, (d) receipt control, (e) data processing and management, and (f) periodic status reporting. Although these tasks may be accomplished manually, it was felt that, in order to produce a high quality survey, a computerized system was necessary.

RTI's experience suggests that the daily operations of a survey are best monitored by a combination of automated and manual processes. The entire second follow-up survey was monitored by such a support system from time of first newsletter mailing until all data had been collected and an edited machine-readable data base was established. Figure E.3 shows the data sources that were inputs to the computerized survey support system, the products that were its outputs, and the interrelations among the various data sets. As can be seen, the logic underlying this control system was that all operations should pass through a central computer-controlled process prior to any further action and that the central process should verify, edit, validate, and schedule activities using the totality of information available at any given time. Table E.3 lists the basic operations performed by the survey control system.

The survey control system was designed to respond to the occurrence of events and to monitor the sequence and relationship of events as they occur over time. Monitored events included mailout of materials (newsletters, questionnaires, transcript requests), nonrespondent sample selection, execution of telephone interviews, receipt control, and data entry and editing. In general, this system maintained data that reflected the current survey status

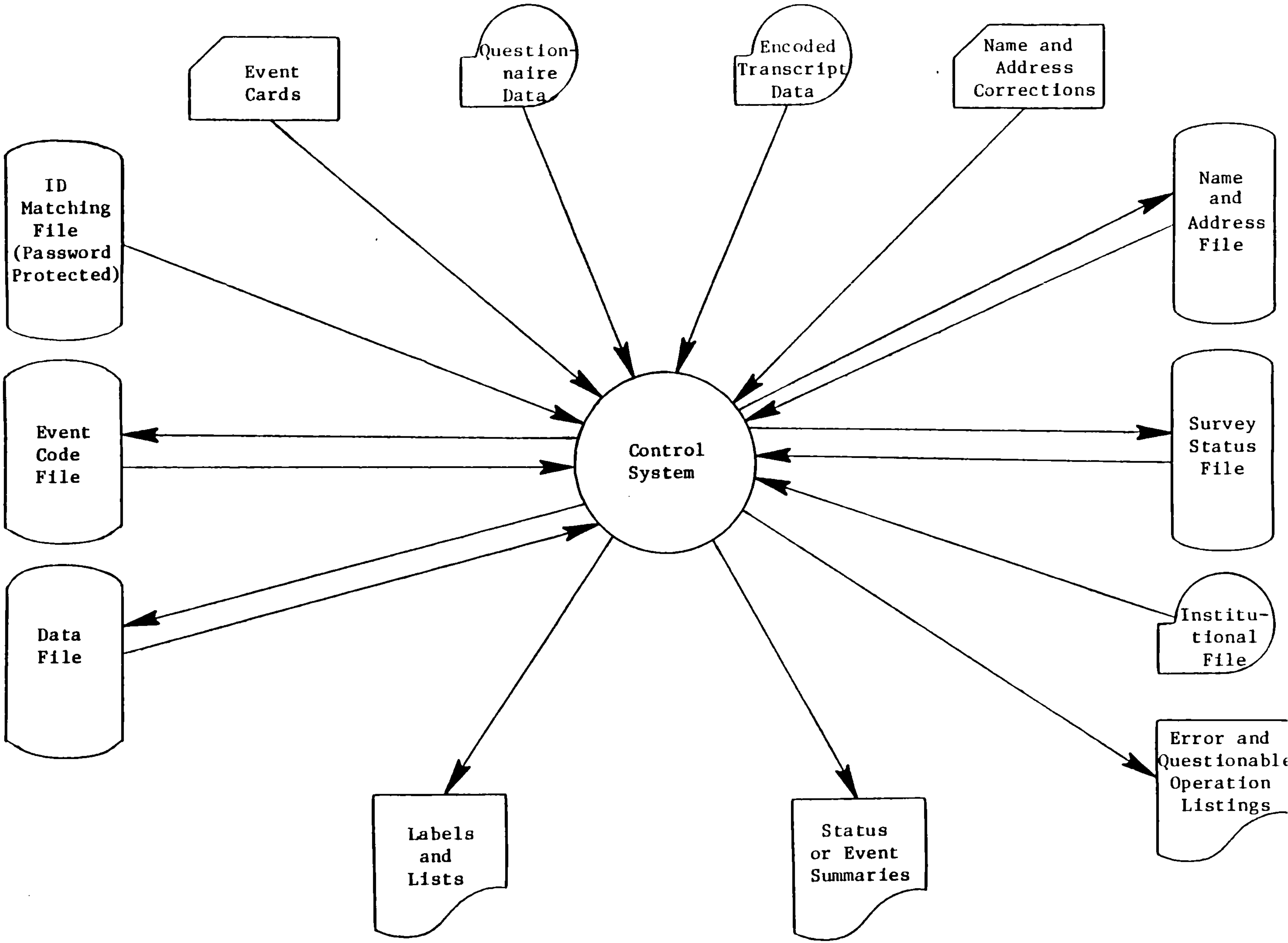


Figure E.3 Study data set elements in relation to survey control system.

Table E.3

BASIC OPERATIONS PERFORMED BY SURVEY CONTROL SYSTEM

Operation	Input Data Sets	Output Operations
1. Generate Student Mailing Labels	Name and Address File ID Matching File Survey Status File	Print List and Mailing Labels Update Survey Status File Update Event Code File
2. Process Events	Event Cards ID Matching File Survey Status File	Print Event Errors (Not Processed) Print Questionable Events Update Survey Status File Update Event Code File
3. Update Addresses	Name and Address Correction Cards ID Matching File Survey Status File	Print Attempted Illegal Update (Not Processed) Update Name and Address File Update Survey Status File Update Event Code File
4. Sample for Telephone Interview	Name and Address File ID Matching File Survey Status File	Print Lists and Labels Update Survey Status File Update Event Code File
5. Produce Status Summaries	Survey Status File	Print Summary
6. Produce Event Summaries	Event Code File	Print Summary
7. Edit and Enter Questionnaire Data	Questionnaire Data ID Matching File Survey Status File Data File	Print Attempted Illegal Entry (Not Processed) Print Fail Edit Listing Enter Edited Data in Data File Update Survey Status File Update Event Code File
8. Generate Labels, Records of Disclosure, and Lists for Transcript Request	Data File Institutional File ID Matching File	Print School Label Print by Institution List of Students Attending Print by Institution Associated Student Names on Records of Disclosure Update Survey Status File Update Event Code File

(continued)

Table E.3

(continued)

Operation	Input Data Sets	Output Operations
9. Edit and Enter Transcript Data	Encoded Transcript Data ID Matching File Survey Status File Data File	Print Attempted Illegal Entry (Not Processed) Print Fail Edit Listing Enter Edited Data in Data File Update Survey Status File Update Event Code File
10. Add Institutional Data	Institutional File Data File ID Matching File	Match Institutional Data and Enter in Data File

of each sample member (survey status file), as well as useful historical information describing the sample member's sequencing through each defined survey event (event code file). In addition to monitoring the survey activities, the system could provide, on request, statistics on each of the operational categories determined by the events that had been defined (e.g., current response status).

To be effective, the system required timely input for each event that took place; therefore, as any of the events occurred for an individual, the event was manually coded and entered into the computerized survey control system on an event card containing the individual's unique ID number. Some events were computer-generated (e.g., selection of an individual for telephone interview); in such cases, the occurrence of the event was entered directly into the system at the time it was generated. In any case, since this processing was performed on a flow basis, the events were reflected in the system within a few days of their occurrence. The specific events in this survey are listed in Table E.4. For each event a specific set of procedures was applied and the survey status file was appropriately updated to allow quantitative monitoring of events and accurate processing of subsequent events.

As events occurred, a card file of pending events was accumulated. A computer program processed the event cards to update the survey status vector of any individual involved in the survey. An individual's status vector contained each of the 19 elements defined in Table E.5. There were 6,050 such vectors, one for each sample member.

In addition to providing a complete and accurate record of survey progress, the support system allowed rigorous quality control of data receipt and management activities through a series of internal checks. Such checks involved accessing an individual's current status prior to performing various data entry operations, which considerably reduced the likelihood of entering data for the wrong individual and virtually eliminated the problem of duplicate record entry. Event records consisted of student ID, check digit, and

Table E.4

MONITORED SECOND FOLLOW-UP SURVEY EVENTS

Event Number	Description	Condition Generating Event
1[a/]	Presurvey Address Update	Address update processed prior to newsletter mailing
2[a/]	First Wave Newsletter Mailed	Label generated for student for first wave newsletter mailing
3[a/]	Delayed Newsletter Mailed	Newsletter tracing update made, leading to label generation, and newsletter previously not mailed
4[a/]	Newsletter Remailed	Newsletter tracing update made, leading to label generation, and newsletter previously mailed
5	Undeliverable Newsletter	Newsletter returned undeliverable
6	Return of Newsletter Postcard	Newsletter postcard returned, no indication of "deceased"
7	Newsletter Indicates Deceased	Newsletter package or postcard returned with indication that sample member is deceased
8[a/]	Newsletter Address Update, P.O.	Newsletter address update processed containing code indicating update is Post Office supplied
9[a/]	Newsletter Address Update, Tracing	Newsletter address update processed containing code indicating update is from tracing
10[a/]	Newsletter Address Update, Student	Newsletter address update processed containing code indicating update is from returned postcard
11[a/]	First Questionnaire Mailed	Label generated for first questionnaire mailout
12	First Questionnaire Undeliverable	First questionnaire mailout returned undeliverable
13	Return of First Questionnaire	First questionnaire returned, no indication of "deceased"
14	First Mailing Indicates Deceased	First questionnaire mailout returned, with indication that sample member is deceased
15	Check Cashed	Notification received that check has been cashed

(continued)

Table E.4

(continued)

Event Number	Description	Condition Generating Event
16[a/]	Second Questionnaire Mailed	Label generated for second questionnaire mailout
17	Second Questionnaire Undeliverable	Second questionnaire mailout returned undeliverable
18	Return of Second Questionnaire	Second questionnaire returned, no indication of "deceased"
19	Second Mailing Indicates Deceased	Second questionnaire mailout returned with indication that sample member is deceased
20[a/]	Data Collection Address Update, P.O.	Data collection update processed containing code indicating update is post office supplied
21[a/]	Data Collection Address Update, Tracing	Data collection update processed containing code indicating update is from tracing
22	Questionnaire Fails Edit	Returned questionnaire fails edit
23	Fail Edit Resolved	Previous fail edit questionnaire resolved
24[a/]	Sampled for Phone F.U.	Label generated for telephone interview
25	Phone Information From Secondary Source Only	Telephone interviewing completed, only secondary source data obtained
26	Phone Information From Sample Member	Telephone interviewing complete, sample member contacted and information obtained from him/her
27	Phone Information Indicates Deceased	Telephone interviewing terminated, sample member reported dead
28[a/]	First Questionnaire Mailing Data Entry	Questionnaire data entered with code indicating first mailing
29[a/]	Second Questionnaire Mailing Data Entry	Questionnaire data entered with code indicating second mailing
30[a/]	Telephone Interview Data Entry	Questionnaire data entered with code indicating telephone
31[a/]	Unidentifiable PSE Institution	Questionnaire data indicate PSE institution cannot be identified

(continued)

Table E.4

(continued)

Event Number	Description	Condition Generating Event
32[a/]	Student Refuses Transcript Access	Questionnaire data indicate identifiable PSE institution and transcript refusal
33[a/]	Student Does Not Refuse Transcript Access	Questionnaire data indicate identifiable PSE institution and no refusal
34[a/]	Transcript Request Mailed	Student name printed on a student listing form
35	Transcript Received	Student transcript data received
36[a/]	Transcript Data Entry	Transcript Data Entered onto Data File
37	Special Tracing Update During Newsletter Operation	Special purpose tracing update required during newsletter tracing operations
38	Terminal Fail Edit (Refusal)	Questionnaire returned indicating complete refusal of sample member to
39[a/]	Remailing of Questionnaire	Previously undeliverable questionnaire remailed after obtaining tracing update
40	Data Collected During Tracing	Tracing operations unable to locate individual but obtained information from secondary source by telephone
41[a/]	Data Entry, Remailing	Questionnaire data entered with code indicating remailing
42[a/]	Data Entry, Combined Form	Questionnaire data entered with code indicating combined form
43[a/]	Data Entry, Tracing Only	Questionnaire data entered with code indicating secondary source data collected during tracing

[a/] These events were generated by computer.

Table E.5

ELEMENTS OF A SAMPLE MEMBER'S SURVEY STATUS VECTOR

Element Number	Status Code Name	Valid Codes
1	Labeling Code	0 = UB, Form A 1 = UB, Form B 2 = CS, Form A 3 = CS, Form B
2	Second Follow-Up Eligibility	0 = Not Eligible 1 = Eligible, No Address 2 = Eligible, With Address
3	Newsletter Mailing	0 = Not Mailed 1 = Mailed, First Wave 2 = Delayed Mail 3 = Remailed
4	Newsletter Response	0 = None 1 = Returned Undeliverable 2 = Post Card Returned 3 = Dead
5	Newsletter Address Update	0 = None 1 = P.O. Supplied Forwarding Address 2 = Tracing Update 3 = Student Supplied Update
6	Questionnaire Mailout 1	0 = Not Mailed 1 = Mailed
7	M1 Response	0 = None 1 = Returned Undeliverable 2 = Questionnaire Returned 3 = Dead
8	Check Cashed	0 = No 1 = Yes
9	Questionnaire Mailout 2	0 = Not Mailed 1 = Mailed 2 = Remailed Undeliverable
10	M2 or Remail Response	0 = None 1 = Returned Undeliverable 2 = Questionnaire Returned 3 = Dead

(continued)

Table E.5

(continued)

Element Number	Status Code Name	Valid Codes
11	Data Collection Address Update	0 = None 1 = P.O. Supplied Forwarding Address 2 = Tracing Update 3 = Data Collected During Tracing
12	Fail Edit Code	0 = No Fail Edit Problems 1 = Fail Edit 2 = Edit Resolution 3 = Terminal Fail Edit (Refusal)
13	Telephone Sample Selection	0 = No Selected 1 = Selected
14	Telephone Contact	0 = No Information 1 = Secondary Source Information 2 = Primary Source Information 3 = Dead
15	Data Entry	0 = None 1 = M1 2 = M2 3 = Remailing 4 = T 5 = Combined 6 = Tracing Only
16	Transcript Permission	0 = No PSE or Indeterminate 1 = PSE Not Identified 2 = PSE Identified, Refusal 3 = PSE Identified, No Refusal
17	Transcript Request Mailed	0 = No 1 = Yes
18	Transcript Received	0 = No 1 = Yes 2 = Unknown at Institution
19	Transcript Data Entry	0 = No 1 = Yes

event number (1-43). In processing any event record, certain checks were performed. As a result of the internal checks, one of three mutually exclusive events occurred:

(1) Change of status vector elements, S(2) through S(19),[7]

(2) Printing of ID number, event number, and status vectors for the individual, with no further action taken on the event, or

(3) Printing of an error message with no further action taken on the event.

These actions represented three error conditions: no error suspected, error suspected or unusual circumstance, and definite error, respectively. The specific checks performed are given in Table E.6.

V. PREMACHINE (MANUAL) PROCESSING OF RETURNED DATA

The steps of preliminary processing of returned data were directed toward ensuring quality data preparation from the returned questionnaires. Most of this preliminary processing of returned questionnaires and transcripts was performed manually. Second follow-up questionnaires and transcripts returned by mail were checked-in by identification numbers on the day they were received, batched for document control purposes, and forwarded to the manual coding/editing staff.

A. Questionnaire Processing

An initial step in premachine processing consisted of a manual scan-edit to see if each returned questionnaire contained a minimum set of key item data. This procedure was completed as soon as possible after receipt (and event-coding of that receipt) and involved only a subset of the questionnaire items with particular emphasis on skip pattern failure and items addressing important study constructs. A specification of these edit rules is provided in Table E.7. Questionnaires that failed the key-item edit were separately batched and routed to the telephone follow-up staff for contact with the respondents to resolve omissions or inconsistencies related to these items, prior to additional coding/editing.

The initial scan-edit (and resolution, if required) was followed by a more comprehensive edit and coding of all items of the questionnaire. Stray pencil marks were erased and intended responses were distinguished from crossed-out responses by colored pencil. Standard alphabetic error codes were used for major item indeterminacy problems, as indicated in Table E.8. Other manual coding was minimal due to the predominantly self-coding nature of the questionnaire, and involved simply assigning Federal Interagency Committee on Education (FICE) or other unique codes to the current or last attended PSE institutions reported by respondents. Specifications for the second stage of questionnaire editing/coding are provided in Table E.9. It should be noted that imputations were performed in this level of edit.

---

7 Status vector element S(1) was not changed once it was assigned an initial value.

Table E.6

EVENT PROCESSING STEPS FOR SECOND FOLLOW-UP STUDY

<u>Event 1: Pre-Survey Address Update</u>

```
IF(S(2).EQ.0), THEN ERROR;
OTHERWISE, IF(S(2).EQ.2), THEN PRINT;
           OTHERWISE, S(2) ← 2.
```

<u>Event 2: First Wave Newsletter Mailing</u>

```
IF((S(2).LE.1).OR.(S(3).NE.0)), THEN ERROR;
OTHERWISE, S(3) ← 1.
```

<u>Event 3: Delayed Newsletter Mailing</u>

```
IF((S(2).NE.1).OR.(S(5).NE.2)), THEN ERROR;
OTHERWISE, IF(S(3).NE.0), THEN ERROR;
           OTHERWISE, S(3) ← 2.
```

<u>Event 4: Remailed Newsletter</u>

```
IF((S(3).EQ.0).OR.(S(5).NE.2)), THEN ERROR;
OTHERWISE, IF((S(3).EQ.2).OR.(S(3).EQ.3).OR.(S(4).GE.2)), THEN PRINT;
           OTHERWISE, S(3) ← 3.
```

<u>Event 5: Undeliverable Newsletter</u>

```
IF(S(3).EQ.0), THEN ERROR;
IF(S(3).EQ.3), THEN PRINT;
IF((S(3).EQ.1).OR.(S(3).EQ.2),
     THEN IF(S(4).NE.0), THEN ERROR;
          OTHERWISE, S(4) ← 1.
```

<u>Event 6: Returned Newsletter Postcard</u>

```
IF(S(3).EQ.0), THEN ERROR;
IF((S(3).EQ.1).OR.(S(3).EQ.2);
     THEN IF(S(4).NE.0), THEN ERROR;
          OTHERWISE, S(4) ← 2;
IF(S(3).EQ.3),
     THEN IF(S(4).EQ.1), THEN S(4) ← 2,
          OTHERWISE, PRINT.
```

(continued)

Table E.6

(continued)

Event 7: Deceased, Newsletter

```
IF(S(3).EQ.0), THEN ERROR;
IF(S(3).EQ.3), THEN PRINT;
IF((S(3).EQ.1).OR.(S(3).EQ.2)),
     THEN IF(S(4).NE.0)), THEN ERROR;
          OTHERWISE, S(4) ← 3.
```

Event 8: Newsletter Address Update, P.O.

```
IF(S(3).EQ.0), THEN ERROR;
OTHERWISE, IF(S(5).EQ.0), THEN S(5) ← 1;
          OTHERWISE, PRINT.
```

Event 9: Newsletter Address Update, Tracing

```
IF(S(5).LE.1), THEN S(5) ← 2;
OTHERWISE, PRINT.
```

Event 10: Newsletter Address Update, Student

```
IF((S(3).EQ.0).OR.(S(4).NE.0)), THEN ERROR;
OTHERWISE, IF(S(5).EQ.3), THEN PRINT;
          OTHERWISE, S(5) ← 3, AND
                     S(4) ← 2.
```

Event 11: Mailing #1

```
IF((S(6).NE.0).OR.(S(4).EQ.3)), THEN ERROR;
OTHERWISE, IF((S(2).EQ.1).AND.(S(5).EQ.0)), THEN ERROR;
          OTHERWISE, S(6) ← 1.
```

Event 12: M1 Undeliverable

```
IF((S(6).EQ.0).OR.(S(7).NE.0)), THEN ERROR;
OTHERWISE, S(7) ← 1.
```

Event 13: M1 Returned

```
IF((S(6).EQ.0).OR.(S(7).NE.0)), THEN ERROR;
OTHERWISE, S(7) ← 2.
```

Event 14: Deceased, M1

```
IF((S(6).EQ.0).OR.(S(7).NE.0)), THEN ERROR;
OTHERWISE, S(7) ← 3.
```

(continued)

Table E.6

(continued)

Event 15: Check Cashed

```
IF((S(6).EQ.0).AND.(S(9).EQ.0), THEN ERROR;
OTHERWISE, IF(S(8).NE.0), THEN ERROR;
           OTHERWISE, IF(S(6).EQ.0), THEN PRINT;
                      OTHERWISE, S(8) ← 1.
```

Event 16: Mailing #2

```
IF((S(9).NE.0).OR.(S(6).EQ.0).OR.(S(7).GE.1)), THEN ERROR;
OTHERWISE, S(9) ← 1.
```

Event 17: M2 or Remail Undeliverable

```
IF((S(9).EQ.0).OR.(S(10).NE.0)), THEN ERROR;
OTHERWISE, S(10) ← 1.
```

Event 18: M2 or Remail Returned

```
IF((S(9).EQ.0).OR.(S(10).NE.0)), THEN ERROR;
OTHERWISE, S(10) ← 2.
```

Event 19: Deceased, M2 or Remail

```
IF((S(9).EQ.0).OR.(S(10.NE.0)), THEN ERROR;
OTHERWISE, S(10) ← 3.
```

Event 20: Data Collection Address Update, P.O.

```
IF((S(6).EQ.0).AND.(S(9).EQ.0)), THEN ERROR;
OTHERWISE, IF(S(11).EQ.0), THEN S(11) ← 1,
           OTHERWISE, PRINT.
```

Event 21: Data Collection Address Update, Tracing

```
IF((S(11).LE.1), THEN S(11) ← 2;
           OTHERWISE, PRINT.
```

Event 22: Fail Edit

```
IF((S(7).NE.2).AND.(S(10).NE.2).AND.(S(11).NE.3).AND.((S(14).EQ.0)
               .OR.(S(14).EQ.3))), THEN ERROR;
OTHERWISE, IF(S(12).NE.0), THEN ERROR;
           OTHERWISE, S(12) ← 1.
```

(continued)

Table E.6

(continued)

<u>Event 23: Edit Resolution</u>

```
IF((S(12).EQ.0).OR.(S(12).EQ.3)), THEN ERROR;
IF(S(12).EQ.2), THEN PRINT;
OTHERWISE, S(12) ← 2.
```

<u>Event 24: Sampled for Telephone Interview</u>

```
IF((S(7).GE.2).OR.(S(10).GE.2).OR.(S(11).EQ.3).OR.(S(4).EQ.3)), THEN ERROR;
OTHERWISE, IF(S(13).GT.0), THEN ERROR;
           OTHERWISE, S(13) ← 1.
```

<u>Event 25: Phone Information, Secondary Source</u>

```
IF((S(13).NE.1).OR.(S(14).NE.0)), THEN ERROR;
OTHERWISE, S(14) ← 1.
```

<u>Event 26: Phone Information, Primary Source</u>

```
IF((S(13).NE.1).OR.(S(14).NE.0)), THEN ERROR;
OTHERWISE, S(14) ← 2.
```

<u>Event 27: Deceased, Telephone</u>

```
IF(S(13).NE.1).OR.(S(14).NE.0)), THEN ERROR;
OTHERWISE, S(14) ← 3.
```

<u>Event 28: Data Entry, M1</u>

```
IF((S(7).NE.2).OR.(S(15).NE.0).OR.S(12).EQ.3)), THEN ERROR;
OTHERWISE, IF(S(12).EQ.1), THEN PRINT;
           OTHERWISE, S(15) ← 1.
```

<u>Event 29: Data Entry, M2</u>

```
IF(S(9).NE.1).OR.(S(10).NE.2).OR.(S(15).NE.0).OR.(S(12).EQ.3)), THEN ERROR;
OTHERWISE, IF(S(12).EQ.1), THEN PRINT;
           OTHERWISE, S(15) ← 2.
```

<u>Event 30: Data Entry, Telephone</u>

```
IF((S(14).EQ.0).OR.(S(14).EQ.3).OR.(S(15).NE.0)), THEN ERROR;
OTHERWISE, IF(S(12).EQ.1), THEN PRINT;
           OTHERWISE, S(15) ← 4.
```

(continued)

Table E.6

(continued)

Event 31: Unidentifiable PSE Institution

```
IF((S(15).EQ.0).OR.(S(16).NE.0)), THEN ERROR;
OTHERWISE, S(16) ← 1.
```

Event 32: Transcript Access Refusal

```
IF((S(15).EQ.0).OR.(S(16).NE.0)), THEN ERROR;
OTHERWISE, S(16) ← 2.
```

Event 33: No Transcript Access Refusal

```
IF((S(15).EQ.0).OR.(S(16).NE.0)), THEN ERROR;
OTHERWISE, S(16) ← 3.
```

Event 34: Transcript Request Mailed

```
IF((S(16).NE.3).OR.(S(17).NE.0)), THEN ERROR;
OTHERWISE, S(17) ← 1.
```

Event 35: Transcript Received

```
IF((S(17).NE.1).OR.(S(18).NE.0)), THEN ERROR;
OTHERWISE, S(18) ← 1.
```

Event 36: Data Entry, Transcript

```
IF((S(18).NE.1).OR.(S(19).NE.0), THEN ERROR;
OTHERWISE, S(19) ← 1.
```

Event 37: Special Tracing Update

```
IF(S(5).LE.1), THEN ERROR;
OTHERWISE, UPDATE (NO CHANGE TO S(5)).
```

Event 38: Terminal Fail Edit (Refusal)

```
IF((S(7).NE.2).AND.(S(10).NE.2).AND.(S(14).NE.0)), THEN ERROR;
OTHERWISE, IF(S(12).NE.0); THEN ERROR;
           OTHERWISE S(12) ← 3.
```

Event 39: Questionnaire Remailing

```
IF((S(7).NE.1).OR.(S(11).NE.2)), THEN ERROR;
OTHERWISE, IF(S(9).NE.0), THEN PRINT;
           OTHERWISE, S(9) ← 2.
```

(continued)

Table E.6

(continued)

<u>Event 40: Data Collected During Tracing</u>

```
IF((S(7).NE.1).OR.(S(10).GE.2)), THEN ERROR;
OTHERWISE, IF(S(11).EQ.2), THEN ERROR;
           OTHERWISE, S(11) ← 3.
```

<u>Event 41: Data Entry, Remailing</u>

```
IF((S(9).NE.2).OR.(S(10).NE.2).OR.(S(12).EQ.3).OR.(S(15).NE.0)), THEN ERROR;
OTHERWISE, IF(S(12).EQ.1), THEN PRINT;
           OTHERWISE, S(15) ← 3.
```

<u>Event 42: Data Entry, Combined Form</u>

```
IF((S(7).NE.2).AND.(S(10).NE.2).AND.((S(14).EQ.0).OR.(S(14).EQ.3))), THEN ERROR;
OTHERWISE, IF((S(12).EQ.3).OR.(S(15).NE.0)), THEN ERROR;
           OTHERWISE, IF(S(12).EQ.1), THEN PRINT;
                      OTHERWISE, S(15) ← 5.
```

<u>Event 43: Data Entry, Tracing Only</u>

```
IF((S(7).GE.2).OR.(S(10).GE.2).OR.(S(14).EQ.1).OR.(S(14).EQ.2)), THEN ERROR;
OTHERWISE, IF(S(15).NE.0), THEN ERROR;
           OTHERWISE, S(15) ← 6.
```

NOTE: One general check, which was applicable to processing all event records (and which is not designated in individual checks specified), was a check of S(2). <u>In processing any even if S(2) = 0, then the ERROR condition was raised</u>.

Table E.7

SPECIFICATIONS FOR KEY ITEM EDIT

1) If response to item 1 is indeterminate, then fail edit.

2) If response to item 1 is "1" and <u>any</u> response is given to items 2 through 10 (excluding the box following Q.10), then fail edit.

3) If response to item 1 is "2" or greater, then

   a) If <u>no</u> response is given to items 2 through 10, then fail edit.

   b) If response to item 2 is indeterminate, then fail edit.

   c) If response to item 2 is "1" and <u>any</u> response is given to items 3 through 10 (excluding the box following Q.10), then fail edit.

   d) If response to item 2 is "2", then

      (i) If <u>no</u> response is given to items 3 through 10, then fail edit.

      (ii) If response to item 4 is indeterminate,[a/] then fail edit.

      (iii) If response to item 6.a, 6.b, 6.c, <u>or</u> 6.d is indeterminate, then fail edit.

      (iv) If response to item 9 is indeterminate, then fail edit.

      (v) If transcript request is denied (i.e., box is checked <u>and</u> name signed in the transcript release statement--following Q.10), then fail edit.

4) If response is indeterminate[b/] to any row of item 11, then fail edit.

5) <u>Form 607-2 only</u>: If response to item 15 is indeterminate,[c/] then fail edit.

NOTE: The term "indeterminate," unless otherwise noted, means no response, illegible response, or multiple response (if applicable).

a/ This item is indeterminate if the PSE institution cannot be uniquely identified (e.g., listing multiple campus school without city).

b/ Note that multiple responses are allowable, as long as one of the responses is not column 1.

c/ Indeterminacy includes the following inconsistent pattern of responses: alternative 1 checked and any other alternative also checked.

Table E.8

STANDARD ERROR CODES USED IN MANUAL CODING AND KEYING

Error Condition	Alphabetic Code	Numeric Code
"Don't Know" response	DK	-5
Refusals	RE	-7
Out-of-range response	OR	-4
Inappropriate multiple response	MR	-2

All questionnaire coding for the second follow-up study, including that performed during manual edit, was done directly on the hard copy (i.e., source document), which had been designed for direct data transcription. Some coding was required for item 4 of the questionnaire, an open-ended item requesting name of school last attended. For each listed PSE institution, a five-digit FICE code with a check digit was entered directly onto the questionnaire (a unique code was used for any sample member school listed in item 4 that was unidentifiable by FICE code). Since a mailing label, which included a unique sample member ID number, had been attached to the questionnaires and to the telephone interview forms, it was not necessary to code the individual's ID number.

Following manual coding and editing, second follow-up questionnaire data, including ID number and batch number,[8] were transformed to machine-readable form by using direct data entry (key-to-tape) programmable terminals. Major advantages of this procedure were higher speed, fewer processing steps, and lower transcript error rate. A programmed format, specifically tailored for the survey questionnaire, appeared on a display screen as the terminal operator keyed the hard copy document. This allowed the operator to make easy corrections of keystroke errors and also to detect, at data entry, any errors that might possibly be corrected by review of the hard copy questionnaire. Error rates are typically lower than with keypunch equipment because operators can correct keystroke errors normally left for the verifier when keypunches are used. Also, 10 percent quality control checks were routinely performed on each batch of data keyed by each operator to maintain a keystroke error rates of less than .5 percent.[9]

8 Keying of batch number facilitated retrieval of hard copy documents (maintained in batch number order) when subsequent error resolution was required.

9 Since this routine RTI quality control process is implemented by sampling the output of each involved key operator for each keyed batch, particularly error-prone batches were identified and completely rekeyed.

Table E.9

SECOND STAGE MANUAL EDIT AND CODING RULES

Item Number	Edit Check or Coding Problem	Error Resolution/Coding
1	Multiple Response	Enter MR in right-hand margin beside item.
2	Multiple Response	Enter MR in right-hand margin beside item.
3	Multiple Response	If multiple response exists, mark through all but numerically greatest (lowermost) response (e.g., if 1 and 6 are circled, mark through 1).
4	a) General	Considerable care in edit and code required. List of PSE institutions and associated codes--with check digits--provided. Code appropriate FICE or other unique code with check digit in right-hand margin beside question. If there is <u>any</u> uncertainty, coding supervisor should be consulted.
	b) Unresolvable or Inappropriate Response	Code OR in right-hand margin beside question.
5	a) General	Round total amounts to nearest dollar value.
	b) Out-of-Range	Any dollar amount greater than $15,000 should be considered out-of-range. That amount should be marked through and OR entered in right-hand margin beside the out-of-range response.
6	a) General	Edit this question separately for each line.
	b) Multiple Response	If multiple response exists on a given line, mark through all but numerically smallest (leftmost) response (e.g., if 1 and 5 are both circled on a line, mark through 5).
7	None	No edit required.
8	Multiple Response	Enter MR in right-hand margin beside item.

(continued)

Item Number	Edit Check or Coding Problem	Error Resolution/Coding
9	a) General	If sub items 2, 3, or 7 are circled, resolution is needed from the coding supervisor. The intent of the question is to obtain information about educational or vocational licensing or certification. Response indicating drivers license, hunting license, or military promotions should not be honored. Where possible, responses to item 7 should be forced into one of the other categories, with an appropriate indication and the response to 7 marked out.
	b) Multiple Response	If "1" is circled and any other number is also circled, mark through the circle around 1.
10	a) General	Edit this question separately for each line.
	b) Multiple Response	If multiple responses exist on any line, then mark through all but the numerically smallest (leftmost) response (e.g., if 1 and 4 are both circled, mark through the 4).
Transcript Request	Code Only	No standard codes are to be used. Code in right-hand margin as follows: 1, if box is checked and name is signed; 0, otherwise.
11	a) General	Edit this question separately for each line.
	b) Multiple Response	If both column 1 and some other column are circled on a given line, then mark through the circle around 1.
	c) Inconsistent Response	1) If "6" is circled on any line and "2", "3", "4", or "5" is circled in some subsequent line (i.e., further down the page), contact coding supervisor. 2) If "2" is circled on any line and "3", "4", "5", or "6" is circled on any previous line (i.e., further up the page), contact coding supervisor.

(continued)

Table E.9

(continued)

Item Number	Edit Check or Coding Problem	Error Resolution/Coding
12	Multiple Response	If multiple response exists in either column, mark through all but the numerically greatest (lowermost) response in the column (e.g., if both 1 and 3 are circled in column 1, mark through the 1 in that column).
13	Multiple Response	Code MR in right-hand margin beside item.
14	Multiple Response	If multiple response exists, mark through all responses except the response nearest to "3". In the event that two responses are equidistant from "3", then mark through the numerically greatest (lowermost) response (e.g., if both 2 and 4 are circled, mark through 2).
15 (Form 607-2)	None	No edit required.
15 (Form 607-1) and 16 (Form 607-2)	a) General	Edit this question separately for each line.
	b) Multiple Response	If multiple response exists on any line, mark through all but the numerically largest (rightmost) response (e.g., if both 1 and 3 are circled, mark through 1).

NOTE: DK and RE codes were applicable to all questions and are not included in these specifications; however, questions requiring edit by line require special consideration for DK and RE coding. If the DK or RE apply to the entire question, then each line is coded appropriately; otherwise, only the applicable line is so coded.

The keying resulted in the creation of a raw data file with transformation of all alphabetic codes to numeric codes. Transformation of error codes is also shown above in Table E.8. The terminals were programmed to accept a specific range of values of most of the data and specific field widths for all data. Specifications for these checks are given in Table E.10.

B. Transcript Processing

The sequence of activities involved in processing student transcripts received from PSE institutions was similar to that described above for questionnaire returns; however, considerably more manual coding and editing were required. Briefly, the transcript coding operation involved abstracting and coding specific items of information from the large and diverse set of PSE transcripts for subsequent keying by the data entry staff. A standard Transcript Coding Form (TCF) and Coding Manual (see Appendix F) was developed for this purpose on the basis of early transcript returns.[10] The manual, which was dynamically revised during the manual coding operation, provided instructions for abstracting important items of information from the diverse transcript formats that were received. This information was manually coded onto a TCF prior to keying.

In addition to overall PSE performance data and academic attendance history, courses appearing on the individual's transcript had to be coded into several content areas, and various performance indices computed within each area. Thus, in contrast to the questionnaire coding, the transcript coding was not a straightforward clerical operation, but one that demanded careful attention and judgment of the coder. Coders for this operation were carefully selected and daily debriefing sessions were held by project staff during the first week of the coding operation to ensure a common base for coding decisions. Also, a quality control check of a sample of each coder's work was conducted weekly by a project staff member.[11]

Prior to initiation of data collection in the second follow-up, it was anticipated that several educational performance measures could be obtained or computed from the information recorded on each student's transcript, including: total number of credits attempted, total number of credits earned, proportion of attempted credits earned, and grade point average. Further, it was expected that each of these indices could be obtained/computed on the basis of the students' overall PSE record, as well as within major field of study and several other curriculum content areas of particular interest to USOE. Because of the uncertainties involved, however, RTI suggested that final specification of such variables and of the procedures for abstracting and coding these data be based on examination of a sample of preliminary transcript returns.

---

10 See also Riccobono, J. A., Wisenbaker, J. M., and Burkheimer, G. J. Description of Procedures Employed in Developing and Implementi‹ ‹ Transcript Processing Manual for the Second Follow-Up Evaluation of the Upward Bound Program. (Report RTI/1605/06 - 03 S). Research Triangle Park, N.C.: Center for Educational Research and Evaluation, Research Triangle Institute, October 1979.

11 These safeguards resulted in reasonably satisfactory estimates of inter-coder reliability for most key transcript variables.

Table E.10

RANGE CHECKS FOR KEYING QUESTIONNAIRES

Item		Valid Codes
Form A	Form B	
ID #[a]	ID #[a]	0001 - 6050
Form Type	Form Type	A, B
Form Seq.	Form Seq.	1, 2, T, R, S
1	1	01 - 04, blank
2	2	01 - 02, blank
3	3	01 - 06, blank
4[a]	4[a]	01001 - 29999, 80000 - 99999, blank
5.i (i-1,15)	5.i (i-1,15)	i[b] or blank
5.16 - 5.18	5.16 - 5.18	00000 - 15000, blank
6.A - 6.E	6.A - 6.E	01 - 05, blank
7.A - 7.D	7.A - 7.D	000 - 999, blank
8	8	01 - 04, blank
9.i (i=1,7)	9.i (i=1,7)	i[b] or blank
10.A - 10.M	10.A - 10.M	01 - 05, blank
TRANS[c]	TRANS[c]	00 - 01
11.A.i -11.V.i (i=1,6)	11.A.i-11.V.i (i=1,6)	i[b] or blank
12.1 - 12.2	12.1 - 12.2	01 - 08, blank
13	13	01 - 03, blank
14	14	01 - 05, blank
NA	15.i - 15.i	i[b] or blank
15.A - 15.H	16.A - 16.H	01 - 04, blank

NOTE: Lead digits are specified; however, these were generated by the program, rather than entered by the key operator.

a/ Excluding Check Digit.

b/ Assumed 2 digits, right justified.

c/ In margin at bottom of page containing Q.10.

Actual examination of transcript returns revealed discrepancies among institutions in transcript content and format that were much more substantial than had been anticipated. In addition to the general problem of major disparity of transcript content and format among the large and diverse set of postsecondary institutions involved in the study, several interrelated problems were identified that precluded or vitiated construction of some of the previously planned indices for particular subgroups of students. These problems stemmed primarily from: (1) the frequency of school transfer behavior exhibited by student sample members together with a lack of complete information concerning the transferred courses listed on the transcript of the school last attended, and (2) the absence, on a large number of transcripts, of specific information that had been expected (e.g., college major). On the other hand, the preliminary examination of transcript returns also suggested several additional variables that could be obtained from information recorded on student transcripts.

After considerable discussion between RTI and USOE staff, decisions were reached regarding the deletion of redefinition of the previously anticipated variables as well as the inclusion of additional variables that seemed to be obtainable from the transcripts that had been collected. Those decisions have been documented separately.[12] Transcript data were keyed in the same manner described above for questionnaires.

VI. MACHINE PROCESSING

All subsequent processing of questionnaire and/or transcript data was computerized and was accomplished on a flow basis. Such processing was undertaken for two purposes: (1) to produce an accurate magnetic transcription of edited second follow-up questionnaires, and (2) to create meaningful analysis variables from these data and data from other external sources.

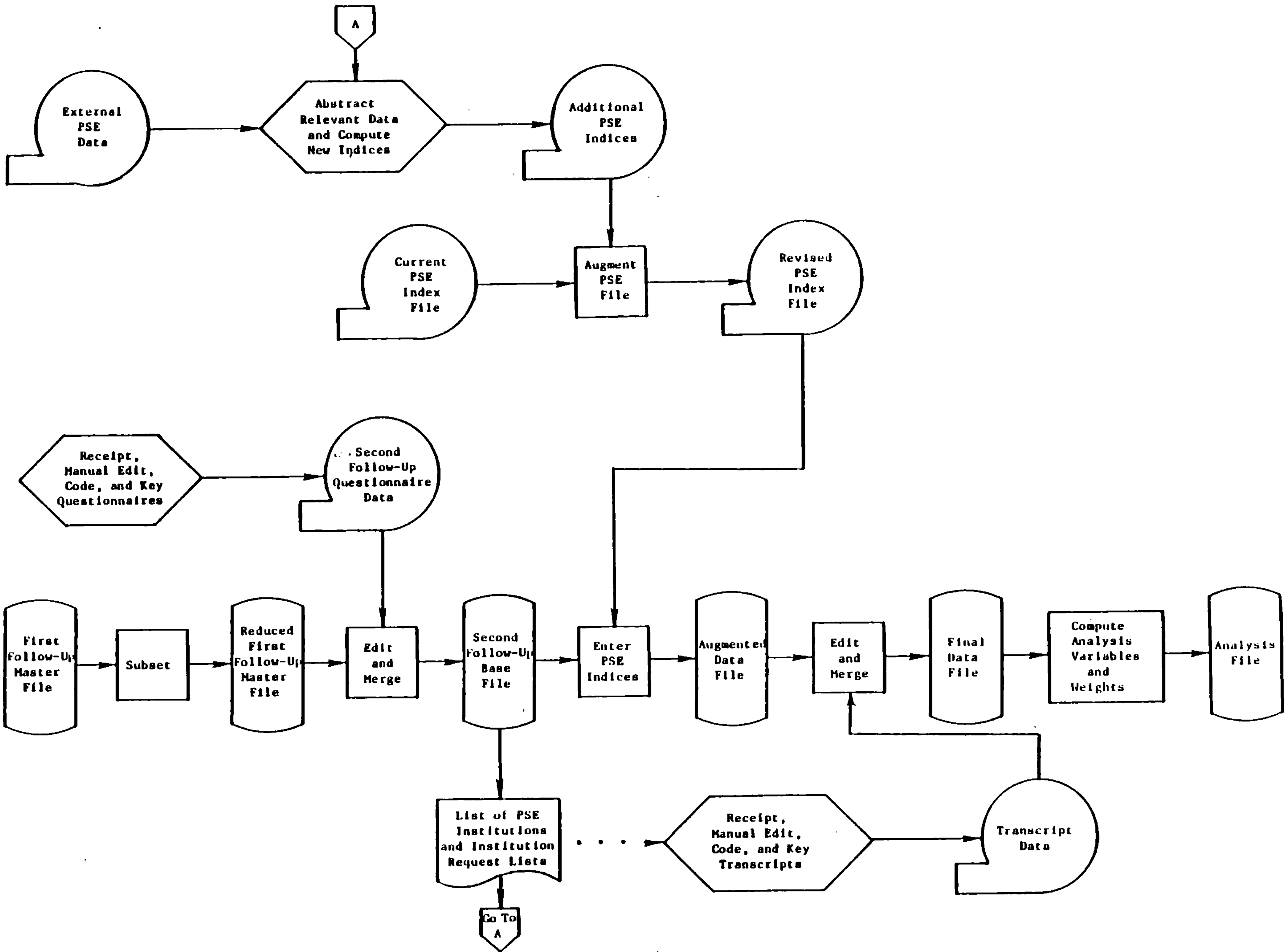
From the sources described in Section II, the second follow-up data base was constructed to be compatible with generalized statistical analysis routines.[13] The final data base consisted of records for each of the original 6,050 sample members containing (1) an encrypted student identification number, (2) selected portions of the extant base-year and first follow-up student data files, (3) a complete history of the individual's response status for the base-year, first follow-up, and second follow-up surveys, (4) a magnetic transcription of student responses to the second follow-up instrument, (5) selected information from an extant postsecondary institution characteristics file, as augmented, (6) analysis variables constructed from data contained in the data sets specified above, (7) analysis variables constructed from information coded from student transcripts, and (8) analysis weights appropriate for the various comparative/descriptive and relational analyses that were performed. The entire data base construction process was coordinated through the survey control system (described above in Section IV). The logical flow of constructing the data base for the second follow-up study is depicted in Figure E.4.

First, information from the previous master file was selected to produce a reduced master file. Following receipt, manual edit and coding, keying, and

---

12 Riccobono, J. A., et al., October 1979 (op. cit.).

13 Burkheimer, G. J., et al., October 1979. (op. cit.)


A
External PSE Data
Abstract Relevant Data and Compute New Indices
Additional PSE Indices
Current PSE Index File
Augment PSE File
Revised PSE Index File
Receipt, Manual Edit, Code, and Key Questionnaires
Second Follow-Up Questionnaire Data
First Follow-Up Master File
Subset
Reduced First Follow-Up Master File
Edit and Merge
Second Follow-Up Base File
Enter PSE Indices
Augmented Data File
Edit and Merge
Final Data File
Compute Analysis Variables and Weights
Analysis File
List of PSE Institutions and Institution Request Lists
Receipt, Manual Edit, Code, and Key Transcripts
Transcript Data
Go To A

transmission of the SFUQ, the data were edited and merged into the data base in a single computer operation. Once this file was completed, lists of PSE institutions attended by sample members and institutional request student listings could be generated. Lists of attended PSE institutions allowed an examination of the Institutional File to determine those institutions for which data augmentation was necessary (due to missing or incomplete institutional indices). Information was then abstracted from available external files to compute the required augmentation indices. These augmentation indices were then merged with the Institutional File to produce a revised institutional index file specific to the attended institutions. Indices for the appropriate institutions were then entered into the student records to produce the augmented student data file.

During the PSE index augmentation process, the institutional transcript request operation was initiated. A single computer operation edited these data, computed appropriate analytic variables, and merged those with the augmented data file to produce the final raw data file. Other analysis variables and weights were then developed from and added to this file.[14] While not shown in Figure E.4, the final step in producing the second follow-up data base was merging the survey control status records and the analysis file. Completion of this operation and subsequent masking or encrypting of identifying information resulted in the data base provided to USOE. A complete specification of the nature of the file developed with a complete file layout has been provided separately[15] and is not considered here. Two major steps in the file construction process (machine editing and analysis variable construction) are detailed in the following subsections.[16]

A. Machine Editing

Once a keyed data tape had been created (either student questionnaire or transcript) the data were merged with the extant Student Master File, following machine edit. The machine edit was undertaken using generalized editing software developed at RTI. Edit programs were tailored for this study following a generalized table-driven procedure supported by existing software, which makes extensive use of machine-readable codebooks.[17]

The additional step of machine edit was undertaken to ensure an accurate magnetic transcription of edited data and to assist in the creation of meaningful and error free analysis variables. Editing consisted of: (1) checking ID number and unique code for institution last attended for accuracy by means of an associated check digit, (2) preliminary data conversions and checks to ensure that all data elements were within prescribed valid code ranges, (3) checks to ensure that data elements within a given source (i.e., SFUQ or transcripts) were internally consistent, and (4) checks to ensure consistency

14 Additional encoding operations were performed during the actual conduct of analyses (e.g., data trimming and combination procedures); however, these operations are described in conjuction with descriptions of the analyses (Appendixes H and I).

15 Burkheimer, G. J., et al., October 1979. (op. cit.)

16 Weight construction is described in Appendix G.

17 These codebooks are tables of information on all data items in the data base (currently collected and previously collected data).

of data from one source with data from other sources (e.g., consistency of SFUQ responses with transcript data and with previously collected data) and special recoding and imputations.

Resolution of errors detected during machine edit of second follow-up data elements took two forms. For key data elements (e.g., PSE institution last or currently attended, extent of UB participation, pattern of PSE attendance) or for cases in which error pattern or frequency suggested coding or keying error, hard copy documents were consulted in an attempt to resolve the error(s). For less important items, and for cases in which resolution from hard copy could not be realized, the involved data elements were appropriately flagged for identification during subsequent analysis.

An error-flagging procedure used by RTI in previous studies (base-year and first follow-up) as well as for the second follow-up study was applied to the data elements. This error-flagging process was accomplished during the editing process. Second follow-up data items (from SFUQ or transcripts) were typically coded in four or more digits. In most cases, the left-most byte or digit contains a consistency digit, and the next leftmost contains a quality digit.[18] The final digits consist of the datum itself and contain either a standard error code or a coded value. The diagram provided below illustrates the layout of a given field.

Consistency Digit	Quality Digit	Datum
1 Digit	1 Digit	2 or more digits

The consistency digit reflects the consistency of the data element with data collected from other sources (i.e., among source consistency). The codes used are:

0 = consistent among sources;
1 = inconsistent among sources.

The quality digit reflects the consistency of the datum within a given instrument (source) as well as an indication of whether the datum was imputed. Codes used are:

0 = Original data;
2 = Data logically imputed;
4 = Data stochastically imputed;
6 = Original data not consistent with internal skip patterns;
8 = Original data internally inconsistent with other data items.

[18] The consistency digit is applicable only to data elements for which consistency checks were available. To conserve space, the consistency digit was only appended to such data elements.

The standard error codes reflect various forms of missing or otherwise indeterminate data. The codes used are:

- -1 = Missing data (item nonresponse);
- -2 = Inappropriate multiple response;
- -3 = Legitimate nonresponse;
- -4 = Out-of-range or illegible response;
- -5 = Don't know responses;
- -6 = Nonresponses inconsistent with routing item;
- -7 = Stated refusal;
- -8 = Items not administered;
- -9 = Instrument nonresponse.

This fairly elaborate coding system allows both an accurate representation of the hard copy and a method whereby data quality could be considered at many different levels when constructing analysis variables.[19]

1. Specific Machine Edit Rules for SFUQ Items[20]

a. ID Verification

ID verification involved check digit verification, range (0001 - 6050) verification, and verification against survey status vectors for data applicability (see Section IV above). Data with faulty IDs or with a suggestion of inappropriate information were not entered into the data file, and generated a list for hard copy resolution.

b. Preliminary Data Conversion and Out-of-Range Checks

Out-of-range checks were analogous to those specified in Subsection V.A (Table E.10). Each item was checked for legitimate error codes and for legitimate non-error values. In addition, the preliminary data conversions specified in Table E.11 were performed. These conversions involved some imputations. If any item was out-of-range, the record was not entered, and an appropriate error message was printed to facilitate hard copy lookup.

c. Check Digit Verification--FICE Code

FICE or other unique institutional codes were verified against the check digit generation algorithm, in addition to the range check specified. Questionnaires failing this check were not entered and generated an appropriate error message to facilitate hard copy look-up.

---

19 The analysis variables themselves do not contain quality or consistency digits, but allow representation by a standard error code (only -1, and -3 are appropriate to analysis variables).

20 It should be noted that the editing rules specified in this appendix were not applied indiscriminately to the raw keyed output file. Rather, extensive preedit examination of the keyed data was implemented, in which resolution of problem records was attempted through examination of hard copy and/or specific *post hoc*. rules. Since these preedit data adjustments were in most cases idiosyncratic, they are not treated in this appendix.

Table E.11

PRELIMINARY DATA CONVERSIONS

Item 1

If Item 1 is blank, then set value of Item 1 to -1.

Item 2

If Item 2 is blank, then set value of Item 2 to -1.

Item 3

If Item 3 is blank, then set value of Item 3 to -1.

Item 4

If Item 4 is blank, then set value of Item 4 to -1.

Item 5

Step 1

If all of Subitems 5.1-5.15 are blank, then set all values to -1; otherwise set any blank value to 0.

Step 2

If all of Subitems 5.16-5.18 are blank, then set all values to -1; otherwise set any blank value to 0.

Item 6

If any of Subitems 6.a through 6.e are blank, then change all blank values to -1.

Item 7

If all of Subitems 7.a-7.d are blank then set blank values to -1; otherwise set blank values to 0.

Item 8

If Item 8 is blank, then set value of Item 8 to -1.

Item 9

(a) If all of Subitems 9.1-9.7 are blank then set all values to -1; otherwise set any blank value to 0.

(b) If Subitem 9.1 has a value of 1 and any of 9.2-9.7 also has a value of 1, then set value of 9.1 to 0.

(continued)

Item 10

If any of Subitems 10.a-10.m is blank, then set the value of that Subitem to -1.

Transcript Box

If Transcript Box entry is blank, then set value of Transcript Box to 0.

Item 11

Item 11 exists in matrix form in the keyed output file, with rows x, $x \leftarrow a,b,\ldots,v$, corresponding to various time periods and with columns y, $y \leftarrow 1,2,\ldots,6$ corresponding to type of enrollment during a time period (see Appendix B). The transformation for this item involves a reduction of item responses to a single response in each time period (row) x.[a/]

Step 1

Convert all blanks to -1.

Step 2

For each x, compute $G_x = \sum_{y=1}^{6}(11.x.y > 0)$.

Step 3

IF $(G_x.EQ.0)$ $R_x \leftarrow 11.x.1$;
IF $(G_x.EQ.1)$ $R_x \leftarrow FUNCT(11.x.y)$;
IF $(G_x.GT.1)$ ERROR;

where FUNCT(11.x.y) is given by:

```
     DO 5 I = 1,6
     IF(11.x.I.EQ.1) GO TO 10.
5 CONTINUE.
     ERROR
10 Rx = I.
```

If the error condition is raised, the record is not processed further, and a printout is generated to facilitate resolution. Otherwise, the vector $\underline{R}$ contains the transformed values for item 11.

(continued)

Table E.11

(continued)

Item 12

(a) If Subitem 12.1 is blank, then set 12.1 to -1.

(b) If Subitem 12.2 is blank, then set 12.2 to -1.

Item 13

If Item 13 is blank, then set Item 13 to -1.

Item 14

If Item 14 is blank, then set Item 14 to -1.

Item 15 (Form B)

If all of Subitems 15.1-15.4 are blank then set all values to -1; otherwise set all blank values to 0.

Item 16 (Form B) or 15 (Form A)

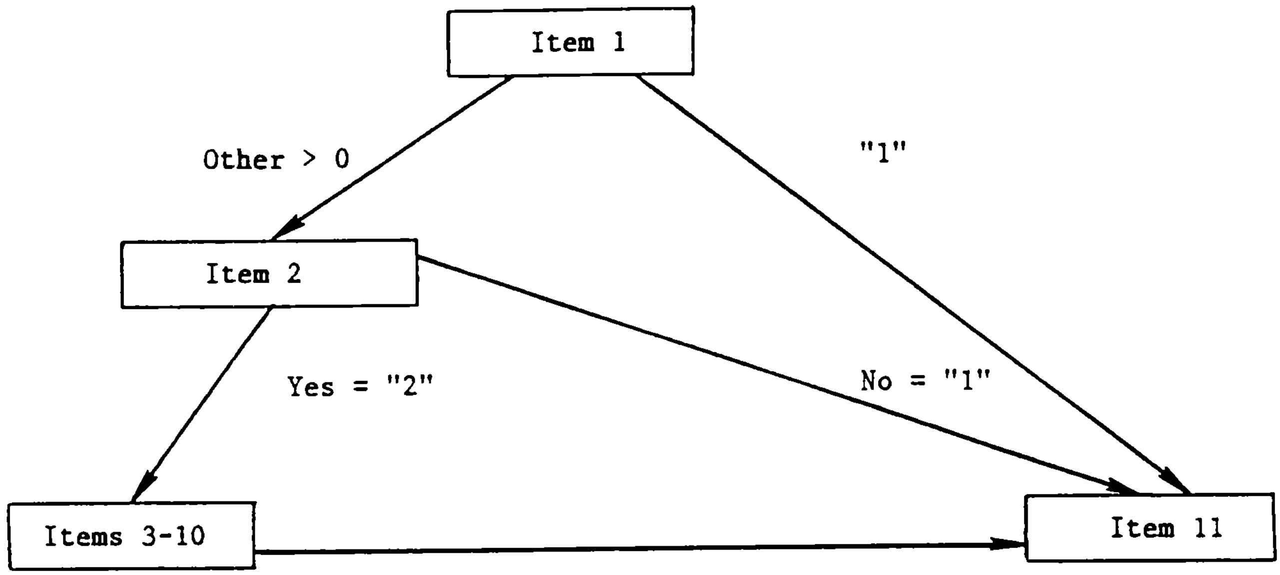
If any of Subitems a-h are blank, then set the value of that subitem to -1.

---

a/ Although this item allowed multiple responses for a given row, preliminary examinations of responses to this item indicated that little information would be lost in such a transformation.

d. Skip Pattern Checks and Procedures

There are two skip patterns (nested) in the SFUQ, that can be represented as follows:



In general, the skip pattern checks that were performed prescribed a special coding with a flag or unique nonresponse code for a routing question that had been ambiguously answered with respect to responses provided to questions within the skip pattern. Also, data for questions within a skip pattern were flagged or specially coded to indicate any inconsistency with the response to the routing question. The flag consisted of a leading six for otherwise legitimate responses, and a negative six code for nonresponse. Appropriately omitted items were also indicated as such by a unique nonresponse code.

Since the patterns were nested, checks were performed from outside-in; that is, a check was made of the outermost skip pattern first. If the check failed, then the inner pattern was not checked, since it was necessarily inconsistent with the outer routing question. There are two ways in which the check of the outermost skip pattern did not fail: (1) the item indicated a skip and all responses within the routing pattern were -1 or (2) the item did not indicate a skip and at least one response other than -1 existed within the routing pattern. In the first case, it was not necessary to perform the innermost check; in the second case, the innermost pattern was checked.

The transcript request blank, which followed Item 10, and which was within both skip patterns., required special treatment. Since this datum was coded only as 0 or 1 (i.e., no nonresponse code is allowed), it was not considered during the skip pattern check. Following a determination of skip pattern failure or non-failure, the item was recoded as indicated in Table E.12.

Table E.12

CODING OF TRANSCRIPT REQUEST

Result of Check	Value of Datum 0	Value of Datum 1
Check Fails	600	601
Check OK (no answer required)	-3	-3
Check OK (answer required)	0	1

For any questionnaire failing a skip pattern check, a printout was generated, identified by ID number, indicating the before and after values of the data elements involved (including the transcript request box value). The general nature of the checks performed and the action taken are given in Figure E.5.

e. <u>Internal Consistency Checks</u>

Internal consistency checks were performed for various combinations of items. In general, the procedure used in performing these checks involved examining logical statements of the form:

[If condition 1 holds for item y, then condition 2 should hold for item z*].

If the logical statement was true, then no action was required; otherwise, each item involved in the check was flagged with a leading 8. This was accomplished by adding a constant to the datum, e.g.

datum $\leftarrow$ datum + $8*10^{(w-2)}$, where w = full field width for the datum (including space reserved for flags).

These consistency checks were not performed until previous edit checks (out-of-range and skip pattern checks) had been performed. The check was <u>not</u> performed if: (1) either field was represented by a value less than 0, or (2) either field was represented by a value equal to or greater than $10^{(w-2)}$, where w is the field width (including space reserved for flags). For any questionnaire failing a consistency check, a printout was generated, identified by ID number, indicating the before and after values of the data elements involved.

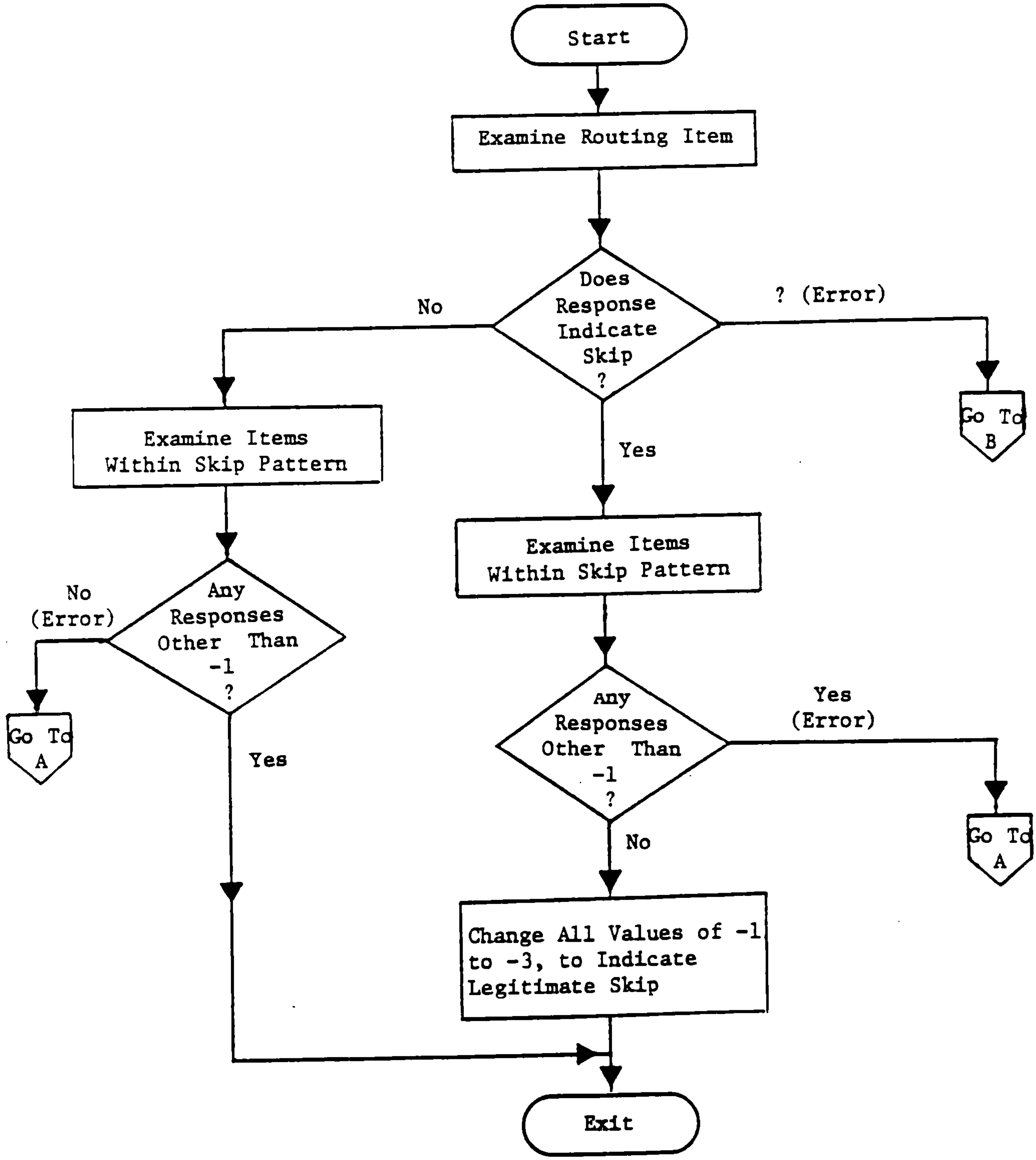


Figure E.5. Edit steps for skip patterns.

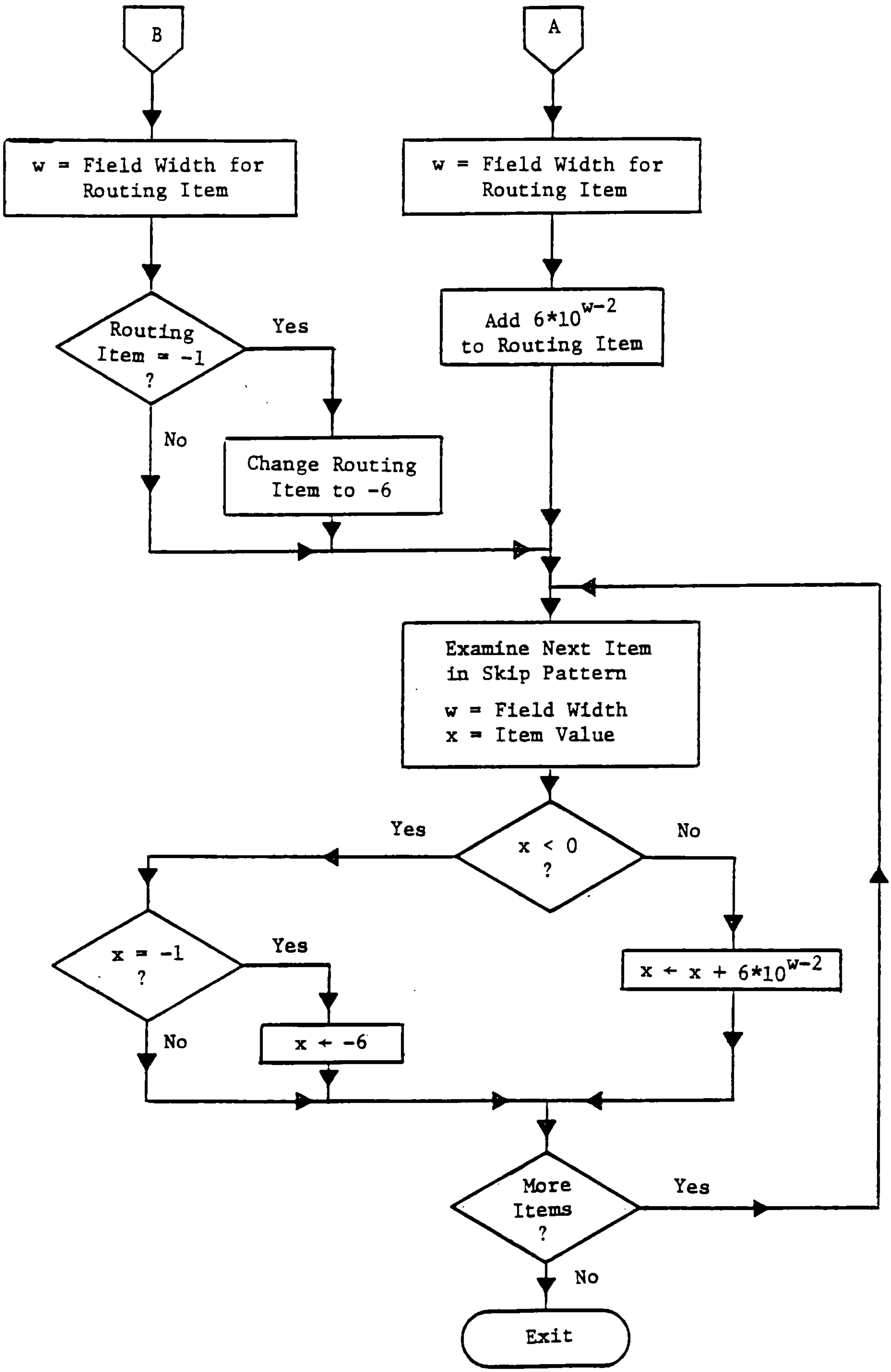


Figure E.5. (Continued)

The specific consistency checks performed were:

(1) Items 1, 2, and 11

If ((Item 1.EQ.1).OR.(Item 2.EQ.1)), then none of 11.x should be greater than 2.

(2) Items 2 and 11

If (Item 2.GE.2), then Subitem 11.v should not be 2.

(3) Item 11

(a) If at least one of 11.x equals 6, let z = the minimum row number in which 11.x = 6; then for any x greater than z, 11.x should not equal 2,3,4, or 5.

(b) If at least one of 11.x equals 3,4, or 5, let z = the minimum row number in which 11.x = 3,4, or 5; then for any x greater than z, 11.x should not equal 2.

(4) Item 12

Let x = response to subitem 12.1; then if (x.GT.0), then 12.2 should be equal to or less than x.

(5) Item 15 (Form B)

If (Subitem 15.1.EQ.1), then each of subitems 15.2-15.3 should be 0.

f. External Consistency Checks

The remaining consistency checks involved checks of SFUQ data with those collected in previous data collection efforts. In general, the procedure used in performing these checks involved examining logical statements of the same form as those indicated above for internal consistency checks (Subsection VI.A.1.e). If the logical statement was true, then no action was necessary; otherwise, each item involved in the check was flagged with a leading 1, this was accomplished by adding a constant to the datum, e.g.

datum $\leftarrow$ datum + $10^{(w-1)}$, where w = full field width for the datum (including space reserved for flags).

These consistency checks were not performed until all other checks (out-of-range, skip pattern, internal consistency) had been performed. The check was not performed if either of the two conditions specified for internal consistency checks (Subsection VI.A.1.e) obtained. For any data failing a consistency check, a printout was generated, identified by ID number, indicating the before and after values of the data involved.

The specific consistency checks performed were:[21]

(1) Items 1 and 2

If ((Item 1.EQ.1).OR.((Item 1.EQ.2).AND.(Item 2.EQ.1)), then (EDHS+.NE.2)

(2) Item 2

If (Item 2.EQ.1), then (PENS.NE.2)

g. Special Recoding and Imputations

(1) Item 5

Subitems can be grouped into three paired sets (X,Y):

X	Y
1-8	16
9-12	17
13-15	18

For each of these three pairings, the recodings specified in Table E.13 were performed, depending on the value of elements in (X,Y).

Table E.13

SPECIFIC RECODINGS FOR ITEM 5

	Value of Y			
Value of X	-1	0	>0	Other
All -1	a	c	a	a
All 0	b	a	e	a
At least one 1	a	d	a	a
Other	a	a	a	a

a) No action

b) Change value of item in Y to 0

c) Change value of all items in X to 0

d) Change value of item in Y to -1

e) Change value of all items in X to -1

[21] Checks also were not performed if FFU variables were greater than 2.

(2) Item 6 (Performed in order given)

(a) If there were any legitimate responses to 6A-6C (e.g., 1-5) and there was also a nonreponse code for 6A-6C (e.g., -1) then the following imputations were performed:

(i) If all legitimate responses were 01, change -1 values to 201;

(ii) If all legitimate responses were 02 or greater, change -1 value to 205;

(iii) Otherwise no imputation.

(b) If there were legitimate (original or imputed) responses to 6A-6C and a nonresponse code for 6D, then the following imputations were performed:

(i) If all responses were 01, change -1 to 401;

(ii) If all responses were 02 or greater, compute the average of the responses (do not count imputation flag, if present), round to nearest integer, and add 400; replace the -1 with this number;

(iii) Otherwise, no imputation.

(c) If there were legitimate (original or imputed) reponses to 6A-6C and a nonresponse code for 6E, then the following imputations were performed:

(i) If all responses were 01, change -1 to 402;

(ii) If all responses were 04 or greater, change -1 to 404;

(iii) If all responses were 02 or 03, change -1 to 401;

(iv) Otherwise, no imputation.

(3) Item 7

(a) If any legitimate response (i.e., a number equal to or greater than zero) existed for 7A-7D, then any nonresponse (-1) was changed to 0 and steps b and c were implemented. If all responses were -1, then no action was taken.

(b) Compute SUM = 7A + 7B + 7C + 7D.

(c) Perform the following recoding:

(i) If SUM $\neq$ 100 then recompute as follows

$$7.i \leftarrow [(7.i \times 100)/SUM] + 2000, \; i \leftarrow A, B, C, D$$

(ii) Otherwise, no action.

(4) Item 10

If there were legitimate responses to any of items 10A-10M and also nonresponse codes for at least one of 10A-10M, the following imputations were performed:

(a) If none of the legitimate responses were 1, then all -1 were changed to 401.

(b) If some legitimate responses were 1, but none were 2, then all -1 were changed to 402.

(c) Otherwise, no imputation.

(5) Item 11

This item existed in a transformed status with a single response for each time period x. If all entries were -1, then no imputation was necessary. Moreover, if the only non-zero entries were 1, then no reasonable imputation could be performed. If none of these conditions were met, then the following imputations were performed.

(a) If some entry was 2, then determine z = max (row number with a 2 entry), and

(i) Change the value of any of rows a, e, i, m, q, or u that was less than z and represented by -1 to 401.

(ii) Change any other row less than z and not one of the rows a, e, i, m, q, or u that was represented by -1 to 402.

(iii) For any row y, where y = c, g, k, o, or s, greater than z that was represented by -1:

A. If response to row y-1 was not missing (i.e., positive) and was identical to that for row y + 1, then change the response for row y to 400 greater than response to row y-1.

B. Otherwise change y to 401.

(iv) For any other row greater than z that was represented by -1 change the value to 401.

(b) If no entry equaled 2, then determine z = min (row number with an entry of 3,4,5, or 6), and

(i) Perform no imputations for rows less than z.

(ii) For any row y, where y = c, g, k, o, or s, greater than z that was represented by -1:

A. If response to row y-1 was not missing (i.e., positive) and was identical to that for row y + 1, then change the response for row y to be 400 greater than the response to row y-1.

B. Otherwise change y to 401.

(iii) For any other row greater than x that was represented by -1, change the response to 401.

(6) Item 12

If response to 12.2 was a legitimate response (i.e., 1-8) and 12.1 was -1, then

$$12.1 \leftarrow 12.2 + 400$$

(7) Item 15 (Form A) and 16 (Form B)

If there was a legitimate response (i.e., 1-4) to any of subitems A-H and also a nonresponse code (i.e., -1) for one or more of subitems A-H, then all -1 were changed to 401.

2. Machine Edit Rules for Transcript Coding Form (TCF) Data

The data from the TCF was used for two major purposes: (a) to augment SFUQ data, and (b) to serve as a basis for construction of analysis variables. USOE had expressed a disinterest in a clean data file of TCF data, so that editing of the TCF file was strictly to support one of the two purposes specified above. Rules for editing and SFUQ augmentation are provided in this subsection. Rules for analysis variable construction are provided in Subsection VI.B.4.b.

a. ID Verification

Student ID verification involved check digit verification as well as a range check. Additionally, the appropriateness of the transcript data was verified through the survey control system. Data with invalid numbers were not entered into the data file, and generated a list for hard copy resolution. Further, the output file should be sorted by student ID # in preparation for merging into the student master file.

b. Preliminary Data Conversion and Out-of-Range Checks

Initial data conversion of blanks, alphabetic data, etc. to appropriate codes were performed prior to merging the TCF data with student questionnaire data; however, these conversions applied only to TCF items 7 and 9. These conversions are specified in Table E.14. Out-of-range checks were performed for the TCF data. This involved checking each item against the legitimate values (after conversion) specified in Table E.15. If any item was out-of-range, the data was not entered, and an appropriate error message was printed to facilitate hard copy lookup.

c. Specific Consistency Checks[22]

These checks were performed for internal consistency of TCF items. Forms failing these checks were not processed, and appropriate error messages were generated.

(1) Check 1

Let N be the number of unique item 10 pages represented for a given record.

IF(TCF8.NE.N), then error.

(2) Check 2

Let K represent the number of non-blank unique 3-tuples of data (standard code, # credits A, # credits B), for a particular entered content area.

IF(K.LE.0), then error.

(3) Check 3

Let each of the unique 3-tuples represent elements of a vector X.

IF(((X_2.EQ.0).OR.(X_2.EQ.blank)).AND.((X_3.EQ.0).OR.(X_3.EQ.blank) .AND.(X_1.NE.40)), then error.

(d) Rules for Augmenting SFUQ Data from TCF Data

Items 7 and 9 of the TCF were used to augment SFUQ data. Following a general edit of the TCF file, these augmentations were implemented

---

22. In processing data for these checks, some anomalies existed for the "page 03" representations of item 10 entries. Particularly, blank "page 03" records were encountered. Two possibilities existed: (1) page number as well as data values blank, (2) valid page number but all data values were blank. This occurred mainly in the event that TCF8 = 0,and the "page 03" record was voided. Under other circumstances, an error condition was raised and no additional processing attempted for that individual.

Table E.14

VALID ENTRY CODES/RANGES FOR TCF DATA

Item	Valid Codes
1[a/]	0001-6050
2	1-6, blank
3	Y, blank
4	Y, blank
5	S, Q
6	Y, N, blank
7.i (i = 1,6)	i or blank
8	0-7
9.A-9.V	01-10, blank
10.CA.1[b/]	1-7 or blank
10.SC.i.j	30, 22, 20, 18, 17, 15, 13, 12, 10, 8 7, 5, 3, 0, 40 or blank
10.A.i.j	0-150 or blank (F4.1)
10.B.i.j	0-150 or blank (F4.1)

a/ Excluding check digit (A-K)

b/ Where i = 1, 2, ..., n
j = 1, 2, ..., 15
n = value of item 8

Table E.15

PRELIMINARY TCF DATA CONVERSIONS

<u>Item 7</u>

(a) If all of Subitems 7.1-7.6 are blank, then set value of 7.1 to 1 <u>and</u> set value of all Subitems 7.2-7.6 to 0.

(b) If Subitem 7.1 has a value of 1 <u>and</u> any of 7.2-7.6 has a value of > 1, then set value of 7.1 to 0.

(c) If Subitem 7.1 has a value of 0 <u>and</u> any of Subitems 7.2-7.6 has a value of > 1, then set the value of the latter Subitem(s) to 1.

<u>Item 9</u>

If any of Subitems 9.a-9.v is blank, then set the value of that Subitem to -1.

---

as specified below:

(1) <u>Item 7 of TCF (item 9 of SFUQ)</u>

```
IF(TCF7.1.EQ.1).OR.(TCF7 otherwise indeterminate), then
take no action;

Otherwise,

    IF(SFUQ9 indeterminate), then

        SFUQ9.1 ← 200
        SFUQ9.7 ← 200, and
        SFUQ9.i ← 200 + TCF7.i, for i ← 2, 3, ..., 6;

    IF(SFUQ9 determinate), then

        IF(SFUQ9.1.NE.0), then SFUQ9.1 ← 200;

        IF((SFUQ9.i.EQ.0).AND.(TCF7.i.EQ.1),
          then SFUQ9.i ← 201, for i ← 2, 3, ..., 6
```

(2) <u>Item 9 of TCF (Item 11 of SFUQ)</u>

```
For i ← a, b, ..., v

IF(TCF9.i is indeterminate), take no action;

Otherwise,
```

(a) If (SFUQ11.i is indeterminate or already imputed), then

(i) IF(TCF9.i < 6), then SFUQ11.i ← TCF9.i,

(ii) IF(TCF9.i ≧ 6), then SFUQ11.i ← 400 + (TCF9.i - 5);

(b) IF(SFUQ11.i is determinate and not previously imputed), then

(i) IF(TCF9.i < 6), then SFUQ11.i ← TCF9.i;

(ii) IF(TCF9.i ≧ 6), then take no action.

B. Constructing Analysis Variables

In constructing analysis variables for the UB second follow-up study, some of the data elements established for the base-year and first follow-up studies were employed. Other variables, however, were constructed from new information, derived primarily from the SFUQ and the STF. Many of the analysis variables, both old and new, were simple recodings of the raw data elements, while others were more complex composites or combinations of raw data elements. Whatever the final make-up of a variable, the primary purpose of analysis variable construction was the operationalization of constructs of interest such as "pattern of UB participation," "educational progress/persistence," or "educational performance" in order to address the research questions statistically.

1. General

Table E.16 provides a list of the conceptual variables (or sets of variables) considered in the UB second follow-up analyses. Three major variable sets may be identified: (a) the criterion variable set (outcome measures), (b) the predictor variable set (process measures), and (c) the concomitant variable set (control measures). The criterion variable set reflects the educational outcomes referenced in the analysis questions (see Chapter 1 of the report). In the main, this set of variables was comprised of responses to the second follow-up questionnaire and information obtained from the student transcripts (or combinations of these variables with those collected in the base-year study or externally). This variable set may be specified as variables 1 through 17 of Table E.16.

The predictor variable set was comprised of one variable; namely, pattern of UB participation (variable 18 in Table E.16) which was derived from base-year and first follow-up information or from item 15 of Form B of the second follow-up questionnaire.[23]

The concomitant variable set was comprised primarily of information collected during the base-year study. The specific variables are age, sex, race,

[23] As noted earlier, Form B of the second follow-up questionnaire was designed for those individuals for whom insufficient previous data had been collected to define pattern/extent of UB participation.

Table E.16

CONCEPTUAL ANALYSIS VARIABLES

1. Characteristics of entered PSE institution (specified by item 3 of SFUQ combined with extant PSE index data).

2. Educational progress (determined from items 1, 2, 9, and 11 of SFUQ, from SMF, and from STF).

3. Educational persistence (determined from items 1, 2, 3, 8, 9, 11, and 12 of SFUQ).

4. High school completion (determined from item 1 of SFUQ second from SMF).

5. Postsecondary entry (determined from item 2 of SFUQ, from SMF, and from STF).

6. Postsecondary graduation (determined from item 9 of SFUQ and from STF).

7. Lag entry to PSE (determined from item 11 of SFUQ and from SMF).

8. Educational performance (determined from STF).

9. Postsecondary Transfer (determined from items 3 and 11 of SFUQ and from the STF).

10. Amount and type of financial aid package (determined from items 5 and 7 of SFUQ).

11. Educational expectations and aspirations (determined from item 12 of SFUQ).

12. Deterents to educational progress (determined from item 15 of Form A of SFUQ or item 16 of Form B).

13. Availability of services and programs (determined from item 10 of SFUQ).

14. Utilization of services and programs (determined from item 10 of SFUQ).

15. Adequacy of financial assistance (determined from item 6 of SFUQ).

16. Current work status (determined from item 13 of SFUQ).

17. Satisfaction with quality of life (determined from item 14 of SFUQ).

18. Pattern of UB participation (determined from item 15 of Form B of SFUQ and from SMF data).

19. Age within grade (determined from SMF).

(continued)

Table E.16

(continued)

20. Race (determined from SMF).

21. Sex (determined from SMF).

22. Grade in Fall 1973 (determined from SMF).

23. Poverty status in Fall 1973 (determined from SMF).

24. Academic risk status Fall 1973 (determined from SMF).

25. UB project characteristics (determined from base year data).

grade level in fall 1973, poverty status, academic risk status, and UB project characteristics[24] (variables 19 through 25 in Table E.16).

Although the taxonomy of variables presented above is a convenient one, it should be recognized that in addressing specific analysis questions the distinction between criterion variables and concomitant variables was not as straightforward as indicated. Specifically, some variables listed in the criterion variable set were used as control variables when considering other criterion variables. For example, educational performance, progress, and persistence are complex outcomes that are interrelated, as well as theoretically related to certain characteristics of the institution attended (e.g., size, level of offering, institutional selectivity), utilization of services and programs, educational aspirations, perceived deterrents, and adequacy of financial aid. While such considerations are not explicit in the analysis questions, they are implicit in any careful analysis of these data.

2. PSE Institutional Indices

Fourteen PSE institutional indices were obtained directly from the Institutional File or from USOE directories of funded UB and SSDS programs.[25] For schools not represented in the Institutional File, these indices were imputed (to the extent possible). Institution identification was determined from responses to item 4 of the SFUQ. The specific PSE institutional indices were developed by USOE and are generally defined in the first ten entries in Table E.17; the next four entries were determined from program directories. Two additional indices (used primarily as cross-classifiers in some of the tabular presentations) related to the type of PSE institution attended were computed, and are shown as the last two entries of Table E.17. The computa[illegible]

[24] Each UB participant was associated with a specific project and comparison students were also associated with a UB project by virtue of their being in a high school served by the project.

[25] Only eight of these indices were actually used in analysis, due to either base rate problems, data indeterminacy problems, or redundancy.

Table E.17

INSTITUTIONAL INDICES OBTAINED OR COMPUTED

Mnemonic	Variable Description
IOFCON	Partial crossing of institutional control (public, private) with level of offering (vocational/technical, two-year college, four-year college)
ISES	Institutional SES (family income characteristics of student body)
ISEL	Relative institutional selectivity (derived from average scores on standardized tests)
ICOST	Composite per-student cost of attending institution
ISIZ	Undergraduate student enrollment
IRACE	Predominant (historical, traditional) race of student body
ISTYP	Specialty training (health related, theological) indicator
DEC10	Decile of freshmen becoming sophomores at institutions
DEC11	Decile of freshmen who graduate at institution
FGPA	Overall GPA for freshmen at institution
UBHOST	UB host status (during 1978) of institution
UBHOSTC	UB host status (during any year from 1974-1978) of institution
SSHOST	SSDS host status (during 1978) of institution
SSHOSTC	SSDS host status (during any year from 1974-1978) of institution
IOFF	Level of offering of institution last attended
ITYP	Highest level of offering of any institution attended

of these latter variables was relatively straightforward and has been separately documented.[26]

3. Project Characteristics

Because of the potential for policy-relevant findings, analyses for both the base-year and first follow-up studies had included investigations of the possible impact of project affiliation or specific characteristics of the associated project; however, none of these previous analyses had been fruitful. In effect, the second follow-up analyses provided an omnibus test for project differences, indicating no unique project effects once student background factors were statistically controlled.

In developing project characteristic variables, it was clear that the variable of simple project association should not be used since previous examination of this variable had proved to be so fruitless. The two project characteristics deveoped were: (1) project emphasis (basic skills approach vs. other approaches) and (2) a measure of aggregate racial characteristics of project participants. These variables were obtained directly from base year sampling files.[27]

4. Student Variables

The conceptual student variables are those listed from 2 to 25 in Table E.16. All the conceptual analysis variables map specific questionnaire items, prior data, or transcript data onto one or more of the analysis questions. As noted above, variables 2 through 17 are directly related to the analysis questions and are considered primarily outcome variables, while variables 18 through 24 may be considered either control or predictor variables.

Some of the student variables or constructs indicated in the analysis questions were relatively straightforward (e.g., current work status, problems experienced), and the items measuring them were also relatively straightforward. Other constructs (e.g., educational performance, progress, and persistence) were much more complex and interrelated, and the attempts to define and measure such constructs reflect this complexity.

a. SMF Student Variables

The variables to be obtained directly from the SMF correspond to variable sets 20-24 in Table E.16. Specific variables used are specified in Table E.17. An additional variable obtained from a combination of SMF entries is a measure of age within grade (AGE GRD). This variable was computed from the SMF variables AGE, reported age in spring 1974, and BGRADE, as defined in Table E.18. The exact nature of this variable is indicated in Table E.19. Other variables were developed using SMF information in combination with SFUQ data, but these variables are discussed below in Subsection VI.B.4.c.

---

26 Burkheimer, G. J., Fuentes, E. J., Riccobono, J. A., and Wisenbaker, J. M., Data Analysis Plan for the Second Follow-up Study of the Upward Bound Program (Report RTI/1605/05-01S). Research Triangle Park, N.C. Center for Educational Research and Evaluation, Research Triangle Institute, March 1979.

27 The variables were not used in any of the analyses documented in this report.

Table E.18

STUDENT VARIABLES OBTAINED FROM THE EXTANT STUDENT MASTER FILE

Mnemonic	Variable Description
RACEC	Race (Black, white, other)
SEXC	Sex (Male, female)
BGRADE	Grade level in fall 1973 (10, 11, 12)
POVSC	Poverty status in 1973 (poverty level, not poverty level)
RISKC	Academic risk status in 1973 (risk, not risk)

Table E.19

PRODUCING AN INDEX OF OVER AGE WITHIN GRADE
(AGE GRD)

	Value of BGRADE			
Value of AGEC	Indeterminate	10	11	12
Indeterminate	-1	-1	-1	-1
14	-1	1	1	1
15	-1	1	1	1
16	-1	1	1	1
17	-1	2	1	1
18	-1	2	2	1
19	-1	2	2	2
20	-1	2	2	2
21	-1	2	2	2

b. STF Student Variables

These variables are those included under conceptual variables 2, 6, and 8 in Table E.16. The variables abstracted from transcripts are shown in Table E.20.[28] Transcript-based variables were constructed only in the case that TCF forms were available and appropriate; otherwise, standard indeterminate data codes were assigned. Rules for construction of these variables are provided below.

(1) Determining the value of MAJOR

IF (TCF2 is blank), then MAJOR ← -1
Otherwise, MAJOR ← TCF2

(2) Determining the value of ACPROB

IF (TCF3 is blank), then ACPROB ← 0
IF (TCF3 is "Y"), then ACPROB ← 1

(3) Determining the value of ACDIS

IF (TCF4 is blank), then ACDIS ← 0
IF (TCF4 is "Y"), then ACDIS ← 1

(4) Determining the value of TGRAD

IF (TCF6 is blank), then TGRAD ← -1
IF (TCF6 is "N"), then TGRAD ← 0
IF (TCF6 is "Y"), then TGRAD ← 1

(5) Special rules for TERNCRED, PCRED.1-PCRED.6, TOTGPA, GPA.1-GPA.6 and TRFIND

IF(TCF8 = 0), then set the variables as follows:

TERNCRD ← 0

TOTGPA ← -1

TRFIND ← 0

PCRED.i ← -1 (for i ← 1, ..., 6)

GPA.i ← -1 (for i ← 1, ..., 6)

Otherwise, these variables were computed on the basis of entries for TCF10.

28 Recall that transcript data were also used to augment student questionnaire data in some cases (Subsection VI.A.2.d).

Table E.20

STUDENT VARIABLES OBTAINED FROM TRANSCRIPTS

Mnemonic	Variable Description
MAJOR	Reported college major
ACPROB	Indicator of having been placed on academic probation
ACDIS	Indicator of having been academically dismissed
TGRAD	Indicator of completion of course of study
TERNCRD	Total earned undergraduate credits (scaled to quarter hours)
PCRED.1	Proportion of total earned credits in education and related fields
PCRED.2	Proportion of total earned credits in humanities
PCRED.3	Proportion of total earned credits in other nonscientific academic fields
PCRED.4	Proportion of total earned credits in social sciences
PCRED.5	Proportion of total earned credits in natural sciences, engineering, and related scientific fields
PCRED.6	Proportion of total earned credits in trade and related nonacademic fields
CAEMP	Content area of emphasis (surrogate measure for college major)
RELCRED	Credits earned relative to normal progression
TOTGPA	Total GPA (at school last attended)
GPA.1	GPA (at school last attended) for courses taken in education and related fields
GPA.2	GPA (at school last attended) for courses taken in humanities
GPA.3	GPA (at school last attended) for courses taken in other nonscientific academic fields
GPA.4	GPA (at school last attended) for courses taken in social sciences
GPA.5	GPA (at school last attended) for courses taken in natural sciences, engineering, and related scientific fields
GPA.6	GPA (at school last attended) for courses taken in trade and related nonacademic fields
TRFIND	PSE transfer indicator

(6) Establishing the Value of TRFIND

For all of seven potential content areas, all 3-tuples of information (standard code, # credits A, # credits B) were examined and if any value of # credits B was other than 0 or blank and was associated with a standard code other than 40, then TRFIND ← 1; otherwise TRFIND ← 0.

(7) Establishing the Value of TERNCRD and PCRED.1-PCRED.6

In computing these variables, a matrix, X, was constructed and initialized to 0 as defined in Table E.21. The columns of the matrix were defined by the various content areas (CA) and specified totals and subtotals. The rows of the table were defined by the nature of the standard codes (1st value in the CA 3-tuple entries) or by the subtotal of several code groupings. The computed elements within rows 1-3 were the sum of credits, from column A (second value of 3-tuple) and column B (third value of 3-tuple), within the appropriate standard code row category and the appropriate column content area (CA)--with appropriate row subtotals and totals. The computed entries of row 4 were the sums (within CA) of the entries in rows 1-3. The computed elements of row 5 were: 1 if any standard code of 40 is encountered within the corresponding column CA, and 0 otherwise. Totals and subtotals of these {0,1} indices were computed as indicated.

Table E.21

ROW AND COLUMN DESIGNATION FOR THE MATRIX X

	Col.1 CA1	Col.2 CA2	Col.3 CA3	Col.4 CA4	Col.5 CA5	Col.6 CA6	Col.7 Sub-total	Col.8 CA7	Col.9 Total
Row 1: 30 codes									
Row 2: 0 codes									
Row 3: 0<codes <30									
Row 4: Total of 1st 3 rows									
Row 5: 40 code indicator									

(a) If($X_{4,9}$.EQ.0), then

If($X_{5,9}$.EQ.0), then error;

Otherwise,
TERNCRED ← -1,
TOTGPA ← -1,
GPA.i ← -1, for i ← 1, ..., 6,
Ignore GPA step, and

If ($X_{5,7}$.EQ.0), then

PCRED.i ← -1, for i ← 1, ..., 6;

Otherwise

PCRED.i ← $X_{5,i}/X_{5,7}$, for i ← 1, ..., 6.

(b) If(($X_{4,9}$.GT.0).AND.(($X_{1,9}$ + $X_{3,9}$).EQ.0)), then

TERNCRED ← 0, and

If($X_{4,7}$.EQ.0), then

PCRED.i ← -1, for i ← 1, ..., 6, and
Go to GPA step.

Otherwise,

PCRED.i ← $X_{4,i}/X_{4,7}$, for i ← 1, ..., 6, and

Go to GPA step.

(c) Otherwise--{(($X_{4,9}$.GT.0).AND.(($X_{1,9}$ + $X_{3,9}$).GT.0))}

TERNCRED ← f{$X_{1,9}$ + $X_{3,9}$}[29], and

If($X_{4,7}$.EQ.0), then

PCRED.i ← -1, for i ← 1, ..., 6, and
Go to GPA step.

Otherwise,

PCRED.i ← $X_{4,i}/X_{4,7}$, for i ← 1, ..., 6, and

Go to GPA step.

(8) <u>Establishing the Value of TOTGPA and GPA.1-GPA.6</u>

In computing these variables, first a matrix, Y, was constructed and initialized to 0 as defined in Table E.22. The columns of

[29] If TCF5 = Q, then f(x) = x; if TCF5 = S, then f(x) = R{1.5 * X}, where R(y) = y rounded to the next highest integer value.

Table E.22

ROW AND COLUMN DESIGNATION FOR THE MATRIX Y

	CA1	CA2	CA3	CA4	CA5	CA6	Sub-Total	CA7	Total
Row 1: 30 code credits									
Row 2: Credits for 0 ± Codes < 30									
Row 3: Credits * Codes for 0 ≦ Codes < 30									
Row 4: 40 code indicator									

matrix Y were identical to those previously described for matrix X. The rows of Y defined particular functions of the standard codes (1st value in the CA 3-tuple entries). Letting the 3-tuples for a given CA be represented by a 3-valued vector, Z, then the CA-specific elements of Row 1 were computed over the j 3-tuples within a content area as

$$\sum_j I_j Z_{2,j} ,$$

$$\text{where } I_j = \begin{cases} 1, \text{ if } Z_{1,j} = 30, \\ 0, \text{ otherwise.} \end{cases}$$

The CA-specific elements of Row 2 were computed as

$$\sum_j H_j Z_{2,j} ,$$

where H_j[30] $= \begin{cases} 1, \text{ if } 0 \leqq Z_{1,j} < 30, \\ 0, \text{ otherwise.} \end{cases}$

The CA-specific elements of Row 3 were computed as:

$$\sum_j H_j Z_{1,j} Z_{2,j} .$$

[30] This is not necessarily 1-I.

The CA-specific elements of Row 4 were computed as:

$$Q\{\sum_j (Z_{1,j} = 40)(Z_{2,j} > 0)\},$$

where $Q(x) = \begin{cases} 0, \text{ if } x = 0, \\ 1, \text{ if } x > 0. \end{cases}$

Finally, the subtotals and totals were taken over the specified CA's, within a row.

Given the matrix, computations were carried out as follows.

(a) TOTGPA

If($(Y_{1,9} + Y_{2,9})$.EQ.0), then

TOTGPA ← -1
GPA.i ← -1, for i ← 1, ..., 6, and
DO NOT go to GPA.i step

Otherwise (i.e., $((Y_{1,9} + Y_{2,9})$.GT.0))

(i) If($(Y_{1,9}/(Y_{1,9} + Y_{2,9}))$.GE.0.30), then

TOTGPA ← -1
GPA.i ← -1, for i ← 1, ..., 6, and
DO NOT go to GPA.i step

(ii) If($((Y_{1,9}/(Y_{1,9} + Y_{2,9}))$.GE.0.10).AND.($(Y_{1,9}/(Y_{1,9} + Y_{2,9})$).LT.0.30))), then

$$\text{TOTGPA} \leftarrow \left[\frac{Y_{3,9} + Y_{1,9}*10}{5*(Y_{2,9} + Y_{1,9})}\right], \text{ and}$$

Go to GPA.i step

(iii) If($(Y_{1,9}/(Y_{1,9} + Y_{2,9}))$.LT.0.10), then

$$\text{TOTGPA} \leftarrow \left(\frac{Y_{3,9}}{5*Y_{2,9}}\right), \text{ and}$$

Go to GPA.i step

(b) GPA.i

If($((Y_{1,8} + Y_{2,8})/(Y_{1,9} + Y_{2,9}))$.GE.0.25), then

GPA.i ← -1, for i ← 1, ..., 6

Otherwise, for $i \leftarrow 1, \ldots, 6$;

(1) If($Y_{2,i}$.EQ.0), then

If(($Y_{1,i}$.EQ.0).AND.($Y_{4,i}$.EQ.0).AND.(($Y_{1,8}$ + $Y_{2,8}$).EQ.0),

then

GPA.i ← -3

Otherwise,

GPA.i ← -1

(2) If(Y_2,i.GT.0), then

(i) If(($Y_{1,i}$/($Y_{1,i}$ + $Y_{2,i}$)).GE.0.30), then

GPA.i ← -1

(ii) If((($Y_{1,i}$/($Y_{1,i}$ + $Y_{2,i}$)).LT.30).AND.(($Y_{1,i}$/($Y_{1,i}$ + $Y_{2,i}$)).GE.0.10)), then

$$\text{GPA.i} \leftarrow \left[\frac{Y_{3,i} + Y_{1,i}*10}{5*(Y_{2,i} + Y_{1,i})} \right]$$

(iii) If(($Y_{1,i}$/($Y_{1,i}$ + $Y_{2,i}$)).LT.0.10), then

$$\text{GPA.i} \leftarrow \left(\frac{Y_{3,i}}{5*Y_{2,i}} \right)$$

(9) <u>Computing the Value of RELCRED</u>

The variable RELCRED was computed using the variable TERNCRED and entries for SFUQ11 as follows:

IF(TERNCRED.EQ.-3), then RELCRED ← -3;

IF(TERNCRED.EQ.-1), then RELCRED ← -1;

Otherwise,

Let

$$N_1 = \sum_{\omega \in A_1} (\text{SFUQ.11.}\omega > 2);$$

$$N_2 = \sum_{\omega \in A_2} (\text{FFUQ.11.}\omega > 2);$$

where

$$A_1 = \{b,c,d,f,g,h,j,k,\ell,n,o,p,r,s,t,v\},$$

and

$$A_2 = \{a,e,i,m,q,u\}.$$

Then,

If ((N1.EQ.0).AND.(N2.EQ.0)), then

RELCRED ← -1;

Otherwise,

$$\text{RELCRED} \leftarrow \frac{\text{TERNCRED}}{N_1 * 15 + N_2 * 7}$$

(10) Computing the Value of CAEMP

The variable CAEMP was developed from empirical rules determined from an examination of the pattern of PCRED.1 through PCRED.6 for those individuals for whom MAJOR was defined. The specific algorithm used was provided by USOE/OED. These rules are given below; however, the variable CAEMP was not used in any analyses reported since it was not sufficiently related to MAJOR.

IF(MAJOR is determinate), then CAEMP ← MAJOR

Otherwise,

IF(PCRED.1 is indeterminate), then CAEMP ← -1;

Otherwise,

Determine a six-valued vector $\underline{Z}$,

$$\text{where } Z_i = \begin{cases} 1, \text{ if (PCRED.i.EQ.}\max\limits_{j}\text{(PCRED.j));} \\ 0, \text{ otherwise;} \end{cases}$$

$$\text{and let } L = \sum_{j=1}^{6} (Z_i) 10^{j-1}$$

CAEMP was determined on the basis of L as follows:

Value of CAEMP	Value of L
1	1, 11, 111, 1011, 11010, or > 110000
2	10
3	100, 110, 1100, 1110, 10100, or 100011

Value of CAEMP	Value of L
4	1000, 1001, 1010, 1101, or 1111
5	10000, 10001, 10010, 10110, 11000, 11001, 11011, 11100, or 11110
6	100000 or 110000

3. SFUQ Student Variables

The variables considered in this section are subsumed under conceptual variables 1-7 and 9-18 of Table E.16. The specific variables that were extracted from the SFUQ (either directly or in combination with SMF data) are listed in Table E.23, with some indication as to their meaning. Details of the operationalization of these variables have been provided separately[31] and are not given here.

The number of student variables listed in Table E.23 is large but was not considered excessive. Moreover, the variables were seen to be sufficiently distinct, conceptually, to stand alone. Further, specific variables were related to specific analysis questions; and, while the variables may be related, it should be recognized that the analysis questions are also related. For these reasons, additional data reduction procedures (principal components approaches or nonmetric multidimensional scaling) seemed unwarranted.

VIII. DATA QUALITY

This section describes the quality of the data that were collected during the course of the second follow-up study. While response rates and sample coverage define the upper limits of data quality, two related problems of data quality that are typically present in questionnaire surveys are: (a) partial data sets, introduced by incomplete, inconsistent, or otherwise indeterminate responses on returned questionnaires, and (b) questionable validity of responses. The various data processing checks and the data quality indicators used in this study (see Subsection VI.A above) provided a ready avenue for investigating data indeterminacy. The receipt of duplicate returns from a subset of individuals (due to overlapping data collection steps) provided a means of examining the reliability of responses. In addition, a limited inter-coder reliability study was performed for the coding process applied to student transcripts.

The section is divided into four subsections, each dealing with a separate data quality problem. Subsection A discusses data obtained from the second follow-up questionnaire and from the base-year data base; Subsection B discusses data derived from an external data file containing institutional indices; Subsection C presents the results of an investigation of follow-up data reliability; and Subsection D presents the results from the reliability study applied to the student transcript data.

---

31 Burkheimer, G. J., *et al*., March 1979 (*op. cit*.)

Table E.23

STUDENT VARIABLES OBTAINED FROM THE SECOND FOLLOW-UP QUESTIONNAIRE

Conceptual Variable	Mnemonic	Variable Description
Pattern of UB Participation	UBPAT1[a/]	Five-valued pattern of UB participation--grades in which participated--
	UBPAT2[a/]	Alternate three-valued pattern of UB participation (nonparticipant, typical, atypical)
Educational Progress/Persistence		
(High School Completion)	EDHS2+[a/]	Progress/persistence through high school graduation, or beyond
(Postsecondary Entry)	PENS2[a/b/]	Progress/persistence through post-secondary entry
(Postsecondary Graduation)	GRAD[c/]	Progress/persistence through post-secondary graduation, licensing, or certification, conditional on school last attended
	PERG[c/]	Progress/persistence attendance ratio related to expressed educational goal
	GSCH[d/]	Progress/persistence through graduate school entry
	NOTRM[c/]	Number of terms enrolled in PSE
	CENRG[c/]	Persistence as reflected in either current enrollment or PSE graduation
	CENRG+[c/]	Persistence as reflected in either current enrollment, PSE graduation, or intent to continue
	FTPT[c/]	Persistence as reflected in intensity (full-time, part-time) of PSE exposure
	SUMSCH[c/]	Persistence as reflected in intensity (proportional summer school enrollments) while in PSE

(continued)

Table E.23

(continued)

Conceptual Variable	Mnemonic	Variable Description
Educational Discontinuity		
(Lag entry to PSE)	LENT2[c]	Delay in entering PSE after high school
(Transfer from two-year to four-year institution	PCHNG[d]	Transfer from non-4-year to 4-year school after PSE entry
Amount and Type of Financial Aid Package	AIDTYP.1-AIDTYP.15[c]	Receipt of 15 specific types of support
	AIDAMT.1-AIDAMT.3[c]	Amount of support received from three general areas: grants, loans, work
	AIDCAT.1-AIDCAT.3[c]	Receipt of support from three general areas
	AIDPAK[c]	Index of pattern of total financial aid package
	AIDAL[c]	Percent of desired total aid to be received from loans and work study
Educational Expectations	EDEXP	Highest education expected
Educational Aspirations	EDASP2	Highest education desired
Deterrents to Educational Progress	EDDET2.A-EDDET2.H	Extent of problems experienced in eight specific areas of potential deterrence
Availability of Services and Programs	GAVAIL.1-GAVAIL.3[c]	Availability during PSE exposure of three general types of services/programs
Utilization of Service and Programs	GPART.1-GPART.3[c]	Extent of use of any of three available types of services/programs, given availability
Adequacy of Financial Assistance	AIDAD1.A-AIDAD1.E[c]	Stated adequacy of five categories of assistance

(continued)

Table E.23

(continued)

Conceptual Variable	Mnemonic	Variable Description
Current Work Status	WORK2	Present employment status
Satisfaction with Quality of Life	SATIS	Expressed extent of satisfaction with current circumstances

a/ Previously collected data from SMF was also used in developing these variables.

b/ This variable is applicable only for those sample members who graduated from high school.

c/ These variables are applicable only for those sample members who have entered some form of PSE.

d/ This variable is applicable only for those sample member who: (1) were twelfth graders in 1973-74 and (2) graduated from a four-year PSE institution.

e/ This variable is applicable only for those sample members who attended a two-year PSE institution.

A. Follow-Up Questionnaire Data Quality

The follow-up questionnaire consisted of 16 items,[32] many of which had multiple parts. Questionnaires were collected by mail, telephone, and in some cases from sources other than students. Table E.24 details the resulting data quality, showing the extent of questionable or indeterminate data in the final data base. Data problems that are internal to the questionnaire (e.g., missing data and internally inconsistent responses), plus those that were "fixed" through logical imputations, are seen to be well within normal limits; in fact, they are typically quite low relative to similar surveys. Given the manual fail edit checks for inconsistency and completeness and subsequent attempts at telephone resolution, the minimal nature of such problems was not surprising.

Data availability of variables from the Student Master File has been previously documented in the base-year[33] and first follow-up[34] reports. The

32 For Form A, only 15 of these items were presented.

33 Burkheimer, G. J., Levinsohn, J. R., Koo, H. P., and French, A. M. Evaluation Study of the Upward Bound Program (Volume IV of A Study of the National Upward Bound and Talent Search Programs). (Four Volumes). Research Triangle Park, N.C.: Center for Educational Research and Evaluation, Research Triangle Institute, April 1976. (ED 131-152)

34 Burkheimer, G. J., French, A. M., Levinsohn, J. R., and Riccobono, J. A. Evaluation Study of the Upward Bound Program: A First Follow-up. Research Triangle Park, N.C.: Center for Educational Research and Evaluation, Research Triangle Institute, September 1977.

Table E.24

DATA QUALITY FOR SECOND FOLLOW-UP QUESTIONNAIRE ITEMS AND VARIABLES

Questionnaire Item	Missing or Otherwise Indeterminate Data	Imputed Data	Inconsistent Data	Total Problems[a/]	
Q1	0	0	37	37	(.9)
Q2	4	0	119	121	(2.9)
Q3	51	0	-	51	(1.2)
Q4	72	0	-	72	(1.7)
Q5.1 → Q5.8	527	0	-	527	(12.4)
Q5.9 → Q5.12	526	0	-	526	(12.3)
Q5.13 → Q5.15	630	0	-	630	(14.8)
Q5.16	786	0	-	786	(18.4)
Q5.17	572	0	-	572	(13.4)
Q5.18	713	0	-	713	(16.7)
Q6.A	414	44	-	458	(10.7)
Q6.B	414	196	-	610	(14.3)
Q6.C	417	157	-	574	(13.5)
Q6.D	425	203	-	628	(14.8)
Q6.E	469	34	-	503	(11.8)
Q7.A	545	11	-	556	(13.1)
Q7.B	545	94	-	639	(15.0)
Q7.C	545	62	-	607	(14.3)
Q7.D	545	115	-	660	(15.5)
Q8	49	0	-	49	(1.1)
Q9.1	61	50	-	111	(2.6)
Q9.2	61	20	-	81	(1.9)
Q9.3	61	4	-	65	(1.5)
Q9.4	61	17	-	78	(1.8)
Q9.5	61	34	-	95	(2.2)
Q9.6	61	44	-	105	(2.4)
Q9.7	61	4	-	65	(1.5)
Q10.A	353	22	-	375	(8.8)

(continued)

Table E.24

(continued)

Questionnaire Item	Missing or Otherwise Indeterminate Data	Imputed Data	Inconsistent Data	Total Problems[a/]	
Q10.B	354	17	-	371	(8.7)
Q10.C	349	23	-	372	(8.7)
Q10.D	351	15	-	366	(8.6)
Q10.E	359	15	-	374	(8.8)
Q10.F	349	16	-	365	(8.6)
Q10.G	375	37	-	412	(9.7)
Q10.H	388	33	-	421	(9.9)
Q10.I	376	32	-	408	(9.6)
Q10.J	374	35	-	409	(9.6)
Q10.K	365	33	-	398	(9.3)
Q10.L	377	34	-	411	(9.6)
Q10.M	349	22	-	371	(8.7)
Q11.A	84	131	-	215	(5.0)
Q11.B	80	12	-	92	(2.2)
Q11.C	81	16	-	97	(2.3)
Q11.D	81	17	-	98	(2.3)
Q11.E	84	138	-	222	(5.2)
Q11.F	77	27	-	104	(2.4)
Q11.G	66	31	-	97	(2.3)
Q11.H	82	29	-	111	(2.6)
Q11.I	82	123	-	205	(4.8)
Q11.J	79	43	-	122	(2.9)
Q11.K	79	40	-	119	(2.8)
Q11.L	81	48	-	129	(3.0)
Q11.M	86	100	-	186	(4.4)
Q11.N	80	56	-	136	(3.2)
Q11.O	81	59	-	140	(3.3)

(continued)

Table E.24

(continued)

Questionnaire Item	Missing or Otherwise Indeterminate Data	Imputed Data	Inconsistent Data	Total Problems[a]	
Q11.P	83	60	-	143	(3.4)
Q11.Q	88	99	-	187	(4.4)
Q11.R	85	64	-	149	(3.5)
Q11.S	84	66	-	150	(3.5)
Q11.T	84	67	-	151	(3.5)
Q11.U	93	103	-	196	(4.6)
Q11.V	88	73	-	161	(3.8)
Q12.1	525	342	178	1,045	(24.5)
Q12.2	920	0	178	1,098	(25.8)
Q13	81	0	-	81	(1.9)
Q14	423	0	-	423	(9.9)
Q15.A	65	0	-	65	(1.5)
Q15.B	66	0	-	66	(1.5)
Q15.C	66	0	-	66	(1.5)
Q15.D	65	0	-	65	(1.5)
Q16.A	424	35	-	459	(10.8)
Q16.B	425	31	-	456	(10.7)
Q16.C	424	26	-	450	(10.6)
Q16.D	427	59	-	486	(11.4)
Q16.E	427	57	-	484	(11.4)
Q16.F	424	44	-	468	(11.0)
Q16.G	425	29	-	454	(10.6)
Q16.H	426	31	-	457	(10.7)

a/ Values given in parentheses represent the percentage of cases with data problems; using as a base the 4,262 cases for which some questionnaire data was available.

availability of the relevant data items among the 4,264 sample members for whom second follow-up questionnaire data were obtained is presented in Table E.25. It is clear that the rates are quite satisfactory, and in only two instances do they fall below 99 percent.[35]

B. Availability of PSE Institutional Data

A second data set consists of the data obtained on PSE institutions. As indicated previously, the PSE indices were obtained from an external data file made available by USOE. Table E.26 presents the rates of available PSE institutional information (for those variables used in analysis) for the 3,022 second follow-up respondents who reported the name of their last attended institution. Among these, the only variable on which information was not completely available was institutional cost which was missing for less than 3 percent of these cases.[36]

C. Availability of Student Transcript Information

The availability of student transcript information is detailed in Table E.27. Of the 21 variables computed from each transcript, 7 were available for fewer than 88 percent of all cases. These 7 included the 6 individual content area GPAs and the classification of major field of study. With respect to the first group of these items, lack of availability should come as no surprise since GPA within an area is defined only for those who had earned credits within that area. In the case of major, some unavailability was expected since not all sample members had been in school long enough to reach the stage at which such a classification of effort was meaningful. In addition, many of the two year schools attended by sample members failed to classify students as to major course of study. These factors resulted in fewer than half the transcripts providing this information.

D. Data Reliability

Due to overlapping data collection activities, duplicate returns were received for 104 individuals.[37] These duplicate returns provided a natural vehicle for an examination of the reliability of the returned data.

The variables have been divided into two sets for reporting purposes. The first set contains those variables which are strictly categorical in nature for which a consistency index was calculated based on agreement of response across questionnaires. This information is set forth in Table E.28. The consistency indices for four of these items exceeded .90 (questions 1, 13, 15C, and 15D). Indices for five other items were between .70 and .90 (questions 3, 6E, 8, 15A, and 15B). Indices for the remaining six items were between .50

[35] Greater indeterminacy of poverty status had been observed in previous studies; the greater indeterminacy of AGE GRADE is due primarily to age indeterminacy, which was observed in previous studies.

[36] In actuality, this figure represents the true availability of data for the first six indices listed; however, for these other indices the missing information was imputed.

[37] The subset of individuals with duplicate returns does not, of course, represent a random or necessarily representative sample of respondents.

Table E.25

AVAILABILITY OF MAJOR CLASSIFICATION VARIABLES FROM STUDENT MASTER FILE

Variable	Percent of Cases Available
SEXC	99.92
RACEC	99.19
POVSC	94.31
RISKC	99.75
GRADE	100.00
AGE|GRADE	93.62

Table E.26

AVAILABILITY OF PSE INSTITUTIONAL INDICES

Variable	Number of Cases with Data	Percent
IOFCON	3,022	100.0
ISES	3,022	100.0
ISEL	3,022	100.0
ICOST	2,942	97.4
ISIZ	3,022	100.0
IRACE	3,022	100.0
UBHOSTC	3,022	100.0
SSHOSTC	3,022	100.0

Table E.27

AVAILABILITY OF TRANSCRIPT INFORMATION

Variable	Number of Cases With Available Data	Percent
MAJOR	977	42.3
ACPROB	2,310	100.0
ACDIS	2,310	100.0
TGRAD	2,191	94.8
TERNCRD	2,169	93.9
PCRED.1-6[a]	2,222	96.2
CAEMP	2,163	93.6
RELCRED	2,168	93.9
TOTGPA	2,037	88.2
GPA.1	693	30.0
GPA.12	1,740	75.3
GPA.13	1,426	61.7
GPA.14	1,547	67.0
GPA.15	1,571	68.0
GPA.16	958	41.5
TRFIND	2,310	100.0

NOTE: Percentages are based on the 2,310 cases for which transcripts were available.

a/ Represents six variables.

Table E.28

RESPONSE CONSISTENCY ASSOCIATED WITH CATEGORICAL
SECOND FOLLOW-UP QUESTIONNAIRE ITEMS

Item	Consistency Index[a/]	Item	Consistency Index
Q1	.96 (99/103)	Q12A	.55 (40/73)
Q3	.75 (50/67)	Q12B	.52 (33/63)
Q6A	.55 (28/51)	Q13	.96 (94/98)
Q6B	.57 (28/49)	Q15A	.83 (35/42)
Q6C	.59 (30/51)	Q15B	.88 (37/42)
Q6D	.57 (27/47)	Q15C	.98 (41/42)
Q6E	.70 (37/53)	Q15D	.98 (41/42)
Q8	.72 (49/68)		

a/ This consistency index is the proportion of cases for which both available responses were identical and, as such, does not include in its computation any cases missing information on a particular variable.

---

and .70 (questions 6A, 6B, 6C, 6D, 12A, and 12B). In general, the items with the highest consistency index are the most objective, while those with the lowest are the most subjective.

The second set contains those variables which are dichotomous or quasi-continuous in nature. The reliability indices which were estimated for these variables are set forth in Table E.29. Estimated reliability indices for eight of the variables were in excess of .80; two reliability estimates were less than .40, and for the remaining variables, reliabilities were between .45 and .75. Given the nature of the data collection process, such results are not unexpected.

E. Reliability of Data Coded From Student Transcripts

While the information on student transcripts was reliable, by definition, the encoding process was expected to introduce some error into results. An important aspect in the application of the transcript coding process was the extent of agreement among raters in encoding the available data according to the rules provided. To examine this facet of transcript coding, a coder reliability study was performed.

In the present instance, reliability refers to the ratio of true to total variance and, as such, corresponds to the notions underlying "classical" reliability theory as expressed by Magnusson.[38] The fundamental problem is

---

[38] Magnusson, David. Test Theory. Reading, Massachusetts: Addison-Wesley Publishing Company, 1966.

Table E.29

RELIABILITY INDICES ASSOCIATED WITH NON-CATEGORICAL SECOND FOLLOW-UP VARIABLES

Variable	Reliability Estimate[a]	Variable	Reliability Estimate
Q2	.90	GAVAIL3	.65
SGCAT	.84	GPART1	.72
LCAT	.61	GPART2	.65
ESCAT	.67	GPART3	.67
Q5.16	.25	Q11.1	.47
Q5.17	.48	Q11.2	.71
Q5.18	.91	Q11.3	.81
Q7A	.51	Q11.4	.66
Q7B	.51	Q11.5	.90
Q7C	.36	Q11.6	-- [b]
Q7D	.57	Q14	.72
Q9A	.86	Q16A	.52
Q9B	.85	Q16B	.64
Q9C	-- [b]	Q16C	.66
Q9D	.45	Q16D	.69
Q9E	.63	Q16E	.55
Q9F	.91	Q16F	.65
Q9G	-- [b]	Q16G	.60
GAVAIL1	.65	Q16H	.46
GAVAIL2	.65		

a/ $\hat{r} = \dfrac{MS_S - MS_E}{MS_S + MS_E}$, where MS_S = mean square for subjects, and MS_E = mean square for error.

b/ Indices for these items were not calculated because no "true" variance was present in the sample of responses obtained.

the identification and estimation of true variance. The approach taken called for a design in which the same five sampled transcripts (representing four unique PSE institutions) were processed by each of the eleven coders used in the actual processing of these data, yielding a fully crossed design for each transcribed variable as indicated in Table E.30.

Within this design, the data considered were the analytic variables constructed from the TCF entries, rather than the raw entries themselves; data were analyzed separately for each analysis variable that was constructed, employing a standard two-way analysis of variance (ANOVA). The general form of the analysis, as it applies to each variable, is given in Table E.31.[39]

Table E.30

CROSSED DESIGN FOR DETERMINING CODER RELIABILITY

	Transcripts				
Coders	1	2	3	4	5
1					
2					
⋮					
11					

Since the only "true" variance in the model defined in Table E.31 is represented by the variance due to transcripts (σ_T^2), an expression for the ratio of this variance to the total variance is needed for an estimate of the reliability of each variable. In considering a similar situation Ebel[40] proposed an estimate for the reliability of ratings:

$$\hat{r} = \frac{MS_T - MS_{TC}}{MS_T + (k-1)\ MS_{TC}} \quad ,$$

for which the mean square terms are those defined in Table E.31 and k is the number of raters involved. Substituting the expected values of the mean squares as set forth in the ANOVA table, this formula simplifies to the ratio of σ_T^2 to $\sigma_T^2 + \sigma_{TC}^2 + \sigma_C^2 + \sigma_{R:TC}^2$.

---

39 Glass, Gene V. & Stanley, Julian C. Statistical Methods in Education and Psychology. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.

40 Ebel, Robert L. Estimation of the Reliability of Ratings. Psychometrika, 1961, 16, 407-424.

Table E.31

ANOVA TABLE

Source of Variance	Degrees of Freedom (df)	Sum of Squares (SS)	Mean Square (MS)	Expected Mean Square (E [MS])
Trancript	4	SS_T	MS_T	$\sigma^2_{R:TC} + \sigma^2_{TC} + 11\sigma^2_T$
Coder	10	SS_C	MS_C	$\sigma^2_{R:TC} + \sigma^2_{TC} + 5\sigma^2_C$
Transcript x Coder	40	SS_{TC}	MS_{TC}	$\sigma^2_{R:TC} + \sigma^2_{TC}$

where $\sigma^2_{R:TC}$ is the variance associated with replications within transcripts and coder cells (not estimable within this design)

σ^2_{TC} is the variance associated with the transcript-coder interaction

σ^2_T is the variance associated with the transcript effect, and

σ^2_C is the variance associated with the coder effect.

The appropriate mean squares and the estimates of the reliability of the continuous and dichotomous variables derived from the transcript coding process are provided in Table E.32. It should be pointed out that these are point estimates based on a sample of only five transcripts and that the standard errors of these estimates are probably quite large as a result. Nonetheless, the estimates provide a starting point for consideration of potential problem areas in implementing the coding process.

In general, the simple dichotomous variables (Indicator of Credit Type, Indicator of Academic Problem, Indicator of Academic Dismissal, Indicator of Completion of Course of Study, and Indicator of Presence of Transfer Credits) show greater estimated reliability than the continuous variables. The reliabilities for the former group of variables range from .77 to 1.00, while those for the latter vary between .10 and .98, and about half of these reliabilities are less than .77. Since the dichotomous variables result from a considerably simpler coding task, this finding comes as no surprise; in fact, it is somewhat surprising that these variables were not more reliably determined (although the reliabilities are generally acceptable).

The major continuous analysis variables, total earned credits and overall GPA, were both measured with reasonably high estimated reliability, .86 and .97 respectively. These constructs show considerable variability in estimated

Table E.32

RESULTS OF THE TRANSCRIPT CODING RELIABILITY STUDY

Description of Variable	MS_T	MS_{TC}	$\hat{r}$
Indicator of Academic Problems[a/]	2.1182	.0182	.91
Indicator of Academic Dismissal[a/]	1.8182	.0182	.90
Indicator of Completion of Course of Study[a/]	2.3636	.0636	.77
Indicator of Presence of Transfer Credits	2.0636	.0436	.81
Indicator of Credit Type[a/]	3.3000	0.0000	1.00
Total Earned Credits	8627.8000	126.1500	.86
Credits in Area 1[b/]	154.5636	8.4636	.61
Credits in Area 2[b/]	3031.5000	21.2800	.93
Credits in Area 3[b/]	751.6818	41.4218	.64
Credits in Area 4[b/]	1553.5000	34.7100	.80
Credits in Area 5[b/]	8456.8453	16.3555	.98
Credits in Area 6[b/]	77.2455	35.4055	.10
Overall Grade Point Average	2.5659	.0070	.97
GPA in Area 1	49.1902	3.3767	.55
GPA in Area 2	8.0426	.0236	.97
GPA in Area 3	43.1286	2.8545	.56
GPA in Area 4	11.1201	.3978	.71
GPA in Area 5	9.7569	.2552	.77
GPA in Area 6	42.4605	3.1186	.53
Indicator of Major Field 1	.4545	.0545	.40
Indicator of Major Field 2	0.0000	0.0000	.00
Indicator of Major Field 3	1.7455	.0355	.81
Indicator of Major Field 4	.0182	.0182	.00
Indicator of Major Field 5	.4545	.0545	.40
Indicator of Major Field 6	.0182	.0182	.00

(continued)

Table E.32

(continued)

Description of Variable	MS_T	MS_{TC}	$\hat{r}$
Categories of Response to Other Certificates, Licenses, Diplomas, or Degrees Earned:			
Cannot Determine	2.0727	.0327	.85
Certificate	0.0000	0.0000	--
License	0.0000	0.0000	--
Vocational/Technical Degree	.1545	.0545	.14
Two Year Academic Degree	1.1091	.0691	.58
Bachelor's Degree	0.0000	0.0000	--
Academic Attendance History:[c/]			
Number Terms Not in School	200.0273	7.3973	.70
Number Terms in High School	0.0000	0.0000	--
Number Terms in Vocational/Technical School	0.0000	0.0000	--
Number Terms in Two Year College	84.5455	.3755	.95
Number Terms in Four Year College	233.5909	2.4209	.90

a/ Binary-valued variables.

b/ Expressed as a proportion of total credits.

c/ Categories of Information Inferred and Information Obtained Directly have been collapsed. Estimated reliabilities for uncollapsed categories appeared to be generally unacceptable, indicating very poor consistency in attributing information to one of these two categories.

reliability, however, when classified by major content area. The proportion of credits earned in categories 2 (Humanities), 4 (Social Sciences), and 5 (Natural Science, Engineering, and related scientific fields) seem to have been determined with relative consistency; reliability estimates range from .80 for category 4 to .98 for category 5. The estimated reliabilities for categories 1 (education and related fields) and 3 (other non-scientific academic fields), however, are marginally acceptable at best (.61 to .64), and the estimated reliability for category 6 (trade and related non-academic fields) is unacceptable (.10).

The reliabilities of computed GPAs within content area display a narrower range (from .53 to .97) but can be divided into two groups. The acceptably reliable variables include GPAs in areas 2, 4, and 5 ($\hat{r}$ = .97, .71, and .77 respectively), while the remaining variables are considerably less reliable and cannot be considered acceptable.

Estimated reliabilities for the categories associated with Major Field were rather low. The highest estimate was obtained for the indicator of Major Field 3 (other non-scientific academic fields) with .81. Reliabilities for Fields 1 (education and related fields) and 5 (natural science, engineering, and related scientific fields) were only .40 each. The estimated zero reliabilities for Fields 4 and 6 are simply indicative of the coding errors associated with the first three fields, and no information was available for Field 2 at all.

Similar problems appeared to have been encountered with respect to the indication of certificates, licenses, diplomas, or degrees earned. The single most reliable category here was that of "cannot determine" with an estimate of .85. The only other two defined reliability estimates were for the categories of Two Year Academic Degree ($\hat{r}$ = .59) and Vocational/Technical Degree or Diploma ($\hat{r}$ = .14). Clearly, the lack of agreement among the raters with respect to these indicators was quite serious in the sample of five transcripts.

Finally, the question of reliability of the Academic Attendance History was addressed through counts of the number of times a particular category of attendance was circled over the 22 terms in question. Initial estimates obtained by maintaining the distinction between directly obtained and "inferred" data were quite poor, indicating the failure of the coders to reliably make this distinction. The figures reported in the table are for the counts in the five categories resulting from collapsing "directly obtained" and "inferred" information. In general, these were acceptable; for the three categories in which reliability was defined, two were at or above .90 (Two-Year College and Four-Year College) and one was .70 (Not in School). Given the intended use of this data to augment student responses, the obtained reliabilities were acceptable.

Appendix F

Transcript Coding Form and Instructions

This appendix contains the materials used in manually coding the student postsecondary academic transcripts that were collected in the institutional request survey. The appendix is organized as follows:

Section I: Instructions for Completing the Transcript Coding Form

Section II: Content Area Code List

Section III: Transcript Coding Form

I. INSTRUCTIONS FOR COMPLETING THE TRANSCRIPT CODING FORM

INSTRUCTIONS FOR COMPLETING THE TRANSCRIPT CODING FORM

General

The Transcript Coding Form (TCF) is the data collection instrument designed for use in this coding task. The items on the TCF pertain to various aspects of postsecondary educational activity and performance. Detailed instructions for completing each TCF item are provided below. Most of the item instructions are illustrated with specific examples. However, these examples are intended as aids to understanding the idea behind coding each item; they should not be treated as if they accounted for all possible instances of the item.

This is an unusually complicated coding task. It is impossible to anticipate every problem or situation that might arise and to specify in advance rules for handling it. In many cases you will have to use your own judgment, based on your general knowledge of what kind of information is desired and why. However, whenever you are unsure about what you should do, please ask. If you encounter a problem or situation that you feel is not covered by the instructions, stop and ask your supervisor for advice on how to proceed.

Preparing to Code a Transcript with the TCF

First, read through the instructions for coding the TCF and be sure you understand what is required for each item. After thoroughly reviewing the TCF and instructions, you should become familiar with the transcript you are going to code. Most (but not all) institutions provide (either as part of the transcript or as a separate information sheet) some descriptive material to aid in understanding their grading and credit system, course numbering system, and the various abbreviations, terms, and symbols which may appear on the student's record. Do not begin coding the transcript on the TCF items until you are familiar with both the TCF and the transcript to be coded.

Instructions for Completing the TCF

This section provides detailed instructions for completing each item on the TCF. Please read these instructions carefully and refer to them frequently when you are coding to be sure that you are assigning the correct code to every TCF item. Use only the codes indicated for each item; certain items allow blanks, others do not. FOLLOW THE INSTRUCTIONS!

Item 1: Student ID Number

This number will be written somewhere on the transcript (typically in felt-tip pen in the upper left-hand or right-hand corner). The number will be a maximum of four digits, followed by a letter, which is a computer check digit. The complete student ID number must be entered in the space beside Item 1; leading zeros and the alpha check digit should be coded. Also enter the student ID number in the top right hand corner of all remaining pages of the TCF. IMPORTANT: Be sure that you transcribe this information accurately.

Item 2: Major Field

Scan the transcript for a space in which major field of study is provided. If the student's major is not provided, leave Item 2 blank. If provided, determine the appropriate content area code (1 through 6) for this field from the "Content Area Code List" provided separately, and enter the code number in the space next to Item 2. For example: If the transcript indicates that the student's major field is (was) psychology, then you should enter "4" (i.e., the Content Area Code for Social Sciences) for Item 2 of the TCF.

Item 3: Academic Probation

Examine the transcript carefully for any indication that the individual was at any time placed on academic probation, or suspended for a finite period of time. (NOTE: Special probation automatically given to transfer students does not apply in this case.) If there is an indication that academic probation or suspension was instituted for the student, enter "Y" in the space next to Item 3; otherwise, leave Item 3 blank.

Item 4: Academic Dismissal

When applicable, some indication of academic dismissal usually (but not always) appears as one of the last entries on a transcript. If there is any indication on the transcript that the student was "dismissed" for academic reasons (even if he was later readmitted), enter "Y" in the space next to Item 4; otherwise, leave Item 4 blank.

Item 5: Credit Type

Determine if the school involved is on a "semester" or "quarter" system, i.e., assigns course credits in terms of semester hours or quarter hours. This is typically explicit in the specification of terms on the transcript (e.g., spring quarter 1976, or spring semester 1976). If the school is on a semester system, enter "S" in the space beside Item 5. If the school is on the quarter system, enter "Q" in the space next to Item 5. If it is impossible to determine whether the school is on a semester hour or quarter hour system, or if the transcript clearly indicates some other system (e.g., "clock hours"), then enter S in the space provided (these hours will all be given 40 codes in Item 10). Be sure to enter either an "S" or a "Q"; do not leave this item blank.

Item 6: Course of Study Completion

Scan the transcript to determine whether or not it indicates that the individual has graduated or completed his/her program of study at the institution. If so, then enter "Y" in the space next to Item 6. If it is obvious that the individual has not completed his course of study (has not earned enough credit to obtain a diploma or certificate of graduation), enter "N" in the space next to Item 6. If it cannot be determined to your satisfaction whether or not the individual has completed his or her major course of study, leave Item 6 blank.

Item 7: Certificates, Licenses, Degrees, or Diplomas Earned

If the information on the transcript indicates that the individual, at any time since high school, has received any license, certificate, etc. (of an academic nature) then circle all of the Item 7 response options 2 through 6 that apply; otherwise, circle option "1" (i.e., cannot determine) only. Note: If Option 1 is circled, then no other alternative should be circled.) Do not leave Item 7 blank.

Item 8: Number of Content Areas Coded

The entry next to Item 8 should be a number from zero to seven, reflecting the number of specific content areas represented on the individual's transcript. NOTE: Zero will be an appropriate entry only in those cases where the transcript (or simply a letter from the registrar) indicates that the student attended the institution for so short a time that no official record of the courses taken or grades received were kept.

In order to complete this item (and to assist, subsequently, in completing Item 10), you should:

(1) Go through the entire transcript and write the appropriate content area code (1 through 6, as indicated on the Content Area Code List) next to each of the courses taken by the individual. Note: It is best, where possible, to key on departmental offering in assigning content area codes to courses listed on the transcript. This is especially useful for courses listed as "Seminar," "General Methods," and the like. If, for example, these courses were offered by the Department of Education, they should be assigned to content area "1." On the other hand, if they were given by the Math Department, they should be assigned to content area "5." Try to assign all courses to one of the six content areas. Search the transcript for clues and use your judgment, but avoid pure guessing.

(2) After completing step (1) above, go through the transcript again and write "7" next to each of the courses that could not be classified into one of the six content areas. Content Area Code "7" has been established specifically for such unclassifiable cases. For example: If the transcript lists a course in "General Methods" and no department or other cues are available to suggest the appropriate content area, assign this course to content area "7."

Special Case: In some cases involving courses transferred from another institution, the transcript may indicate the total number of credits transferred but not the actual courses. In such cases, you should write in "7" on the transcript next to the total number of credits transferred, since it is impossible to identify the actual content area(s) for these transferred courses/credits.

(3) Simply count the number of different content area codes that you have assigned to the courses or credits listed on the transcript, and enter this number in the space provided next to Item 8 of the TCF.

It should be pointed out that the number entered for this item must equal the number of "Content Area Data Sheets" (i.e., the number of TCF Item 10 pages) that are completed for the individual. Thus, it is very important that Item 8 be completed correctly, since this number will be used by the data entry personnel to determine how many pages of Item 10 (Content Area Data Sheet) to expect. Do not leave this item blank.

Item 9: Academic Attendance History

Item 9 should be completed from the information provided on the transcript, to the extent possible. In completing this item, no more than one number on each line is to be circled; however, if the information required for a particular line (or time period) cannot be obtained directly or inferred adequately from the transcript, then do not circle any number on that line.

High School Attendance: Information as to when the individual completed high school can frequently be obtained directly from dates explicitly provided in the transcript. During that period of time, the individual was obviously attending high school, and for that period of time, Alternative 2 should be circled. Inferences may also be made for previous years of high school attended by assuming that the individual was in school during the academic year (i.e., fall, winter, and spring) and not in school during the summers. These assumed or inferred entries should be entered as "6"s or "7"s for the appropriate time periods.

Postsecondary School Attendance: Since the transcripts that you will be processing will be batched by institution, it is only necessary to determine from the separate list of postsecondary institutions that will be provided the type of postsecondary school that the institution represents (i.e., vocational-technical or business; 2-year college; 4-year college). During the time the individual was enrolled in the school from which the transcript was obtained, the appropriate alternative (3, 4, or 5, depending on the type of school) should be checked. Note: For a semester system, the fall semester includes both fall and winter of the year under consideration, and the spring semester includes both the winter and spring of the year under consideration. Thus, if a transcript from a four-year college indicates that the student was enrolled in the spring 1974 semester, you should circle "4" for row d and also circle "4" for row c.

For transfer students, the task becomes somewhat more difficult. In some cases, the transcript will indicate the exact semesters or quarters that the individual attended "other" institutions, prior to transferring to the institution in which he is currently or was last enrolled. In this case, the task will be quite simple. You will have to determine the type of school from the separately provided list and then simply circle the appropriate alternative (3, 4, or 5) for the periods of time in which the individual was at the previous school. More often, however, the transcript will indicate nothing more than that the individual was at some other school during a more broadly defined time period (e.g., from 1974 to 1976). In these cases, you will have to infer the actual times that the individual was in school (and circle the appropriate alternatives (8, 9, or 10) for each time period). As a rule of thumb, it is safe to assume, unless otherwise indicated, that the individual was at the previous school during the entire academic year, and not during the summers of the years under consideration. Additional clues as to how long the individual attended previous schools may be gained from the number of credits that are transferred. If, for example, it was indicated that an individual was at School X from 1974 to 1975, and he is transferring only three credits, it is fairly safe to assume that he was only there during the fall semester or quarter of that year. <u>Special Case</u>: The transcript may indicate that the student attended two different institutions during the same time period (e.g, a two-year college). In such cases, you should indicate his attendance at the higher level institution only.

<u>Item 10</u>: <u>Content Area Code</u>

A separate "Content Area Data Sheet" (i.e., Item 10 page) must be completed for each content area code represented by the courses listed on the individual's transcript. The number of these sheets that will have to be completed for a particular transcript is indicated by the number entered for Item 8 above (the procedure for determining this number was described previously under that item).

The procedure to be followed in completing Item 10 is the same for each content area. Specifically, <u>each</u> Content Area Data Sheet (Item 10 page) should be completed as described below.

First, enter the appropriate content area code (1 through 7) in the space provided next to Item 10. (For example, if "education" courses are listed on the student's transcript--and if they are, then you should have already written "1" next to each of these courses on the transcript (see Item 8 instructions)--enter a "1" in the space beside Item 10.) Next, complete columns A and B of the table for this content area as follows:

Column A (Number of Credit Hours at Last School Attended)

(1) Identify any courses in this content area for which the student has received a passing or satisfactory grade (i.e., "P" or "S") and has earned credits toward a degree but which are not included in the computation of his/her grade point average (i.e., there are no quality points associated with passing the course). Enter the total number of credit hours associated with all courses in this content area that were assigned such a grade in the space (under column A in the table) corresponding to the row designated "Passed." Draw a line through these courses and their corresponding grades and credit hours with a "magic marker," so that this information is not mistakenly included again in the other computations for this table. If no courses in the content area have received such a passing grade, then simply leave this space blank and proceed to the next grade (row) in the table.

(2) Repeat the procedure described in (1) above for the remaining grades A+ through "No Grade or Credit Reported" in the table. That is, identify all courses in the content area with a grade of A+, sum the credit hours associated with these courses, enter this total in the designated space under column A, and mark through the courses involved on the transcript. If there are no courses in this category, then leave this space blank. Repeat this procedure for courses given A, and so on.

Special Cases: Some problems (requiring special handling) may result from the presence of fail, withdrawal, and/or "ambiguous" grades given for courses listed on the student's transcript.

For example:

(a) Failing Grades: In many schools, if an individual receives a failing grade for a particular course (i.e., an "F" or "U"), the number of credit hours associated with that course is not listed or is listed as zero. In such cases, you will have to obtain or infer the number of credit hours attempted from other information on the transcript and, after doing so, record it in the appropriate space on the student's transcript. For example, if the student subsequently repeated the course or obtained a passing grade, then the number of credit hours will be recorded there and can be obtained directly. In other cases, the number of credit hours for the failed course can be deduced from the transcript's reported "total hours attempted," especially if this information is reported for each term. If these or similar approaches fail, then assume the "typical" credit hours assigned to similar courses (e.g., if the student failed English 101 and most other English and liberal arts courses on the transcript are three-credit courses, then assume English 101 is also a three credit-hour course.)

(b) WF (Withdrawal Failing), I (Incomplete), NR (Not Reported), H (Grade Held), or Similar Grades: These grades are to be treated in exactly the same manner as Fs or Us. Credit hours associated with courses receiving such grades in a particular content area should be included with those associated with "failed" courses in the area, and the total figure entered under Column A in the space corresponding to "Standard Code 0" on the Content Area Data Sheet.

(c) W or WP Grades (i.e., Withdrawal "Without" Penalty) and "Audits": Credits listed for courses receiving such grades are not to be entered on the Content Area Data Sheets.

(d) NC (No Credit) Grades: The procedure to be employed for such grades will be based upon other information on the individual's transcript. Specifically, you must ask yourself if the grade (1) carries "nonzero" credits and (2) enters into the student's GPA. If the answer to either one of these questions is yes, then the grade should be considered the same as an F and coded accordingly. However, if the answer to both questions is no, then the grade should be treated the same as a W (i.e., withdrawal without penalty) and excluded from further consideration.

(e) No Grade or Credit Reported: In such cases, you should simply enter "1" in columns A and B (on the appropriate Content Area Data Sheet) in the spaces allotted corresponding to Standard Code "40." Courses taken under a "clock hour" credit system should also be coded in this manner.

Column B (Number of Credit Hours Transferred)

The information in this column of the table should be completed only if the transcript indicates that the student has transferred courses/credits from another postsecondary institution; otherwise, column B should be left completely blank. It is anticiated that transfer information will be present on roughly 30 percent of the transcripts.

Transfer data is recorded on the student's transcript in different ways by different institutions. Some institutions provide a complete list of all transferred courses, by term, with corresponding grades and credit hours; however, most institutions only list the transferred courses and their associated credit hours. In any case, the procedures to be followed in completing column B of the table are the same as those described above for column A, with one major exception: if no grades are provided for the transferred coures, only the first line in column B need be completed. In this case, simply enter the total number of credit hours associated with transferred courses in the content area in the space corresponding to "Passed" and leave the other spaces in column B blank.

<u>NOTE</u>: Transferred credit hours (Column B entries) must be coded in the <u>same</u> credit type used by the last school attended (Column A). If the credit system used by the previous (transfer) school is different from that used by the transcript school, you must recompute the transferred credits--unless they have already been converted--using the following conversion rule:

Quarter Hours = Semester Hours x 1.5

or

Semester Hours = Quarter Hours x .67

II. CONTENT AREA CODE LIST

Education and Related Fields (Code=1)

Education, General
Pre-Elementary Education
Kindergarten Education
Elementary Education, General
Secondary Education, General
Junior High School Education
Higher Education, General
Junior College Education
Community College Education
Continuing Education
Adult Education
Special Education, General
Education of Mentally Retarded
Education of the Gifted
Education of the Culturally Disadvantaged
Speech Correction
Education of the Emotionally Disturbed
Remedial Education
Special Learning Disabilities
Education of the Physically Handicapped

Social Foundations of Education
Philosophy of Education
History of Education
Educational Psychology
Educational Statistics
Educational Research
Educational Testing
Educational Measurement
Educational Evaluation
Educational Counseling
Educational Guidance
Educational Administration
Educational Supervision
Curriculum
Instruction

Reading Education
Art Education
Music Education
Mathematics, Education
Science Education
Physical Education
Safety Education
Driver Education
Health Education
Family Life Education
Distributive Education
Commerce Education
Business Education
Industrial Arts Education
Technical Education
Vocational Education
Education, Other (e.g., Bilingual Teacher)

Home Economics, General
Textiles, Home Economics
Home Management
Family Relations
Dietetics
Foods and Nutrition

Humanities (Code=2)

Foreign Languages, All
Teaching of English
as a Foreign Language
Sanskrit
English, General
English, Literature
Literature, Comparative
Classics

Reading
Linguistics
Semantics
Philology
Phonetics
Rhetoric
Public Address
Speech
Debate

Creative Writing

Philosophy

Theological Profession, General
Biblical Languages
Education, Religious
Religious Studies

History

Black Culture Studies
Afro American Studies
African Studies
American Indian Cultural Studies
Mexican-American Cultural Studies
Asian Studies
India Studies
Islamic Studies
Slavic Studies
Russian Studies
Latin American Studies
Middle Eastern Studies
European Studies, General
American Studies
Pacific Area Studies

Other Nonscientific Academic Fields (Code=3)

Agriculture, General
Agronomy
Crop Management
Field Crops
Soil Management
Soil Conservation
Husbandry, Animal
Wildlife Management
Fish Management
Game Management
Fruit Production
Horticulture
Vegetable Production
Ornamental Horticulture
Floriculture
Farm Management
Agricultural Management
Agricultural Economics
Agricultural Business
Food Technology
Forestry
Natural Resources Management
Forestry Technology
Agriculture Technologies
Range Management

Business Management
Business Administration
Operations Research
Restaurant Management
Hotel Management
Institutional Management
Cafeteria Management
Marketing
Purchasing
Public Utilities
Transportation
Insurance
International Business
Secretarial Studies
Personnel Management
Industrial Relations
Labor Relations

Economics, Business
Consumer Economics
Communication, General
Journalism
Radio
Television
Advertising
Communication Media

Business, General
Commerce, General
Accounting
Business Statistics
Banking
Finance
Investments
Securities

Law, General

Community Services, General
Public Administration
Park Management
Recreation Management
Social Work
Clinical Social Work
Helping Services
Law Enforcement
Corrections
Rehabilitation Services
International Public Service
Public Affairs and Services, Other (e.g., Fire Science)

Library Science, General

Fine Arts, General
Sculpture
Painting
Art
Commercial Art
Drawing
Art History
Art Appreciation
Music, Theory
Music, Performing
Music Composition
Music (Liberal Arts Program)
Music History
Music Appreciation
Musicology
Dramatic Arts
Dance

Ceramics
Weaving
Photography

Social Sciences (Code=4)

Social Sciences, General

Psychology, General
Experimental Psychology (Animal)
Clinical Psychology
Counseling, Psychology
Social Psychology
Psychometrics
Statistics in Psychology
Industrial Psychology
Developmental Psychology
Physiological Psychology
Learning Theory
Child Development

Anthropology
Archeology
Economics
Geography
Government
Political Science
Sociology
Criminology
International Relations
Urban Studies
Demography

Natural Science, Engineering, and Related Scientific Fields (Code=5)

Mathematics, General
Mathematics, Applied
Statistics, Mathematical and Theoretical

Military Science
Naval Science
Aerospace Science

Physical Sciences, General
Physics, General
Molecular Physics
Nuclear Physics
Chemistry, General
Inorganic Chemistry
Organic Chemistry
Physical Chemistry
Analytical Chemistry
Pharmaceutical Chemistry
Astronomy
Astrophysics
Atmospheric Sciences
Meterology
Geology
Geochemistry
Geophysics
Seismology
Earth Sciences, General
Paleontology
Oceanography
Metallurgy

Biology, General
Botany, General
Bacteriology
Plant Pathology
Plant Pharmacology
Plant Physiology
Zoology, General
Pathology, Human
Pathology, Animal
Pharmacology
Physiology, Human
Physiology, Animal
Microbiology
Anatomy
Histology
Biochemistry
Biophysics
Molecular Biology
Cell Physiology
Cell Biology
Cytology
Marine Biology
Biostatistics
Biometrics
Ecology
Entomology
Genetics
Radiobiology
Nutrition, Scientific
Neurosciences
Toxicology
Embryology
Limnology
Icthtyology
Parasitology
Ornithology

Health Professions, General
Health Care Administration
Hospital Administration
Nursing
Dentistry
Dental Specialties
Medicine
Medical Specialties
Occupational Therapy
Optometry
Osteopathic Medicine
Pharmacy
Physical Therapy
Dental Hygiene
Public Health
Medical Record Librarianship
Podiatry
Biomedical Communication
Veterinary Medicine
Speech Pathology
Audiology
Chiropractic
Medical Laboratory Technologies
Dental Technologies
Radiologic Technologies

Environmental Design,
General
Architecture
Interior Design
Landscape Architecture
Urban Architecture
Naval Architecture
Regional Planning
City Planning
Community Planning

Engineering, General
Astronautical Engineering
Aerospace Engineering
Aeronautical Engineering
Agricultural Engineering
Architectural Engineering
Biomedical Engineering
Bioengineering
Chemical Engineering
Petroleum Refining
Petroleum Engineering
Transportation Engineering
Civil Engineering
Construction Engineering
Communication Engineering
Electrical Engineering
Electronics Engineering
Mechanical Engineering
Geological Engineering
Geophysical Engineering
Management, Engineering
Industrial Engineering
Metallurgical Engineering
Materials Engineering
Ceramic Engineering
Textile Engineering
Mining Engineering
Mineral Engineering
Engineering Physics
Nuclear Engineering
Engineering Mechanics
Environmental Engineering
Marine Engineering
Ocean Engineering
Engineering Technologies
Sanitary Engineering

Computer Sciences, General
Information Sciences, General
Information Systems
Data Processing
Computer Programming
System Analysis

Trade and Related Nonacademic Fields (Code=6)

Agricultural Mechanics
Agricultural Production
Gardening
Pet Grooming
Modeling (Apparel and Accessories)
Floral Design
Floristry
Food Distribution
Food Services
Cooking
Realty
Recreation and Tourism
Design (Clothing)
Tailoring
Fashion Design
Air Conditioning Installation and Repair
Appliance Repair
Television Repair
Radio Repair
Auto Repair
Auto Mechanics
Automotive Service, Other
Aircraft Mechanics
Aircraft Maintenance
Blueprint Reading
Business Machine Maintenance
Carpentry

Cabinetmaking
Electricity, Construction
Heavy Equipment Mechanics
Heavy Equipment Maintenance
Masonry
Bricklaying
Concrete Finishing
Paperhanging
Painting and Decorating
Plastering
Plumbing and Pipefitting
Sheetrock Installation
Drywall Installation
Glazing
Roofing
Construction and Maintenance, Other
Diesel Mechanics
Drafting
Cartographics
Electric Motor Repair
Small Engine Repair
Communication Assembly Repair, Installation, Etc.

Industrial Electronics
Fabric Maintenance
Piano Tuning
Instrument Maintenance and Repair
Machine Tool Operations
Welding and Cutting
Tool and Die Making
Smelting
Metalworking, All Other
Metalsmithing
Barbering
Cosmetology
Hair Styling
Molding
Bartending
Meat Cutting
Knitting
Spinning
Textile Fabrication
Textile Design
Leatherworking
Upholstering
Woodworking
Trade and Industrial Occupations, Other
Jewelry
Cinematography
Interior Decoration

III. TRANSCRIPT CODING FORM

Transcript Coding Form

1. Student ID Number ______________________

2. Major Field ______________________

3. Academic Probation ______________________

4. Academic Dismissal ______________________

5. .Credit Type ______________________

6. Course of Study Completion ______________

7. Certificates, licenses, diplomas, or degrees earned:

(Circle all that apply.)

Cannot determine 1

Certificate 2

License 3

1-year, 2-year, or 3-year vocational degree or diploma . . . 4

2-year academic degree 5

4-year or 5-year college Bachelor's degree 6

8. Number of Content Areas Coded __________

Student ID #

9. Academic Attendance History:

	INFORMATION OBTAINED DIRECTLY FROM TRANSCRIPT					INFORMATION INFERRED				
	Not In School	High School	Voc-Tech School	2-Year College	4-Year College	Not In School	High School	Voc-Tech School	2-Year College	4-Year College
a. Summer 1973	1	2	3	4	5	6	7	8	9	10
b. Fall 1973	1	2	3	4	5	6	7	8	9	10
c. Winter 1973-74	1	2	3	4	5	6	7	8	9	10
d. Spring 1974	1	2	3	4	5	6	7	8	9	10
e. Summer 1974	1	2	3	4	5	6	7	8	9	10
f. Fall 1974	1	2	3	4	5	6	7	8	9	10
g. Winter 1974-75	1	2	3	4	5	6	7	8	9	10
h. Spring 1975	1	2	3	4	5	6	7	8	9	10
i. Summer 1975	1	2	3	4	5	6	7	8	9	10
j. Fall 1975	1	2	3	4	5	6	7	8	9	10
k. Winter 1975-76	1	2	3	4	5	6	7	8	9	10
l. Spring 1976	1	2	3	4	5	6	7	8	9	10
m. Summer 1976	1	2	3	4	5	6	7	8	9	10
n. Fall 1976	1	2	3	4	5	6	7	8	9	10
o. Winter 1976-77	1	2	3	4	5	6	7	8	9	10
p. Spring 1977	1	2	3	4	5	6	7	8	9	10
q. Summer 1977	1	2	3	4	5	6	7	8	9	10
r. Fall 1977	1	2	3	4	5	6	7	8	9	10
s. Winter 1977-78	1	2	3	4	5	6	7	8	9	10
t. Spring 1978	1	2	3	4	5	6	7	8	9	10
u. Summer 1978	1	2	3	4	5	6	7	8	9	10
v. Fall 1978	1	2	3	4	5	6	7	8	9	10

Student ID # ______

Content Area Data Sheet

10. Content Area Code ______

Grade or Situation	Standard Code	A. Number of Credit Hours At Last School Attended	B. Number of Credit Hours Transferred
Passed*	30	______	______
A+	22	______	______
A	20	______	______
A-	18	______	______
B+	17	______	______
B	15	______	______
B-	13	______	______
C+	12	______	______
C	10	______	______
C-	8	______	______
D+	7	______	______
D	5	______	______
D-	3	______	______
F or U	0	______	______
No Grade or Credits Reported	40	______	______

*Includes transfers with no grade given and grades of "S" and "P".

Appendix G

Sample Weight Adjustments

Appendix G

Sample Weight Adjustments

I. OVERVIEW

Three conceptually different weight adjustment procedures are discussed in this appendix, all of which were employed to provide more meaningful analyses. The first procedure discussed below is a technique whereby two artificial groups were constructed such that the comparison group would be similar to the UB participant group on relevant student background variables (sex, race, poverty status, and academic risk status, on initial grade cohort distribution, and in total group size. In essence, the procedure forces a greater similarity between the two groups in terms of characteristics that are theoretically related to outcomes but that were not completely controlled by the study design. This adjustment procedure (a form of partial balancing) was used exclusively for the relational analyses reported in Chapter 4.[1]

The second procedure discussed in this appendix involves the adjustment of sample weights for the subsampling procedures used, for instrument non-response, and for item indeterminacy. The need for weight adjustment due to subsampling is rather intuitive and obvious. The adjustment for nonresponse or item indeterminacy was used to correct results for possible bias introduced by the failure to contact some individuals or for the provision of incomplete data by some individuals. Such adjustments involved an apportioning of sampling weights of nonrespondents among the respondents who were most like them. These adjustment procedures were used only for the hypothesis-testing analyses reported in Chapter 3.[2]

The final procedure considered involves the adjustment of sampling weights for those individuals originally selected into the comparison group who subsequently reported participation in the UB program (crossovers). The procedure was used to adjust for the fact that crossovers, by virtue of their representing a much larger population and being selected at considerably lower sampling rates, carried considerably greater sampling weights than those originally sampled as UB participants. The adjustment involved a form of weight smoothing that reduced the discrepancies in the weights between the originally sampled UB participants and the crossovers, while maintaining the distributional properties of the originally sampled UB participants. These adjustment procedures also were used exclusively for the hypothesis-testing analysis reported in Chapter 3.[3]

---

1 Although the variables considered in this adjustment (excepting group size) were also included as predictors in the structural equations, the procedure was implemented to reduce certain technical problems introduced through markedly different sampling weight sums (reflecting similar differences in the populations represented by the two groups) and differential distributions within the groups on the variables of interest.

2 For the exploratory relational analyses, assumptions of randomness of nonresponse were made, and differential sampling and subsampling probabilities were not considered.

3 Assumptions made in weight constructions for the exploratory relational analyses did not consider original sampling weights; therefore, the smoothing procedure was not required.

Discussion of the actual procedures used is provided in subsequent sections of this appendix. The remainder of this introductory section will present certain considerations basic to the use of the techniques.

A. Considerations of Corrections for Group Differences

In cases such as the current one, where moderating variables (individual background variables and initial grade cohort) are related to criterion variables (educational outcomes) and where the moderating variables are differentially distributed within the groups under consideration (former participants and nonparticipants), it is usually sound analytic practice to adjust or correct the criterion variables for the influence of the moderating variables. There are many techniques available to accomplish this adjustment process. Perhaps the most widely used is linear regression, in which the linear relation of moderating variable to criterion variable is removed (e.g., analysis of covariance, partial correlation techniques). A very simple form of this technique is well adapted to quantitative data when one is willing to assume that the majority of the moderator's influence is linear and equivalent for different subgroups.

Another approach to data adjustment is the analysis of variance model. This technique, considered within the framework of general linear models, may be used to examine the "effects" (both main effects and interaction effects) of specific categorical factors after having removed any effect attributed to other categorical factors. Like the linear regression model, relatively strong assumptions about the data must be made (e.g., equality of variances over subgroups), and this technique is best applied to quantitative criterion variable data. It would be difficult to employ a complete analysis of variance model, since the number of cases in some cells of the model would be extremely small. Further, since much of the criterion variable data was stated in proportions, considerable data transformation (i.e., log or arc sine transformations) might be required, and reporting such transformed data often clouds interpretation.

For the specific purpose of the adjustment procedure needed for the relational analyses, neither of these approaches were considered appropriate. Instead, adjustments were based on procedures that are similar to those used by demographers in constructing "adjusted," "standardized," or "corrected" birth and death rates.[4]

B. Considerations of Adjustments for Instrument and Item Nonresponse

For the complex sample design used in the UB evaluation, individuals were selected into the sample with unequal probabilities. Therefore, to inflate the sample to the size of the population being estimated (so that each person represented the correct number of persons in the population), sample weights that were inversely proportional to the probability of selection were applied to individual responses. (See Appendix A for a discussion of the use of sample weights.) In producing these national estimates, some type of adjustment was needed to correct for the failure of individuals to return questionnaires or to return fully completed questionnaires.

---

[4] See Spiegelman, M. Introduction to Demography. Cambridge, Massachusetts: Harvard University Press, 1968.

The two most common techniques for handling indeterminate data are imputation techniques and weight adjustment techniques. There are many specific imputation techniques available, but all of them involve creating data for those individuals who have not responded to a particular item. Some techniques assign the mean value of similar individuals; others assign the values of the one individual most like the nonrespondent; still others involve simple or iterative regression models to predict (according to specified criteria) what the missing data most likely would be. Similarly, there are various available weight adjustment procedures; however, they all involve assignment of the weights of nonrespondents to those who responded to a particular instrument or item.

In general, the method of weight adjustment used in this study involved the proportional distribution of a nonrespondent's sampling weight to that group of responding individuals judged to be most similar to him/her. In terms of estimates of means and proportions obtained, this technique yields results that are equivalent to assigning the nonrespondent an imputed value of the group mean of similar respondents. In terms of standard error computation, the two techniques do not yield equivalent results; larger standard errors are obtained by the weight adjustment method. RTI chose the more conservative weight adjustment procedure.

C. Considerations of Weight Smoothing for Crossovers

During the first follow-up evaluation, a number of cases with conflicting UB participation classification were discovered. Some individuals who were originally selected into the UB participant group indicated that they had never participated in UB, and some individuals who were originally selected into the comparison nonparticipant group reported UB membership (in some cases prior to the original sampling). The first group (previously classified as UB participants) consisted of individuals who were either inappropriately defined in the sampling frame or who did not report their UB participation for some reason (forgot about it or did not wish to report it). Various forms of bias could have entered the results by inclusion of these cases (the type of bias suspected varies with the reason for the conflicting classifications), and these cases were excluded from analysis.

The second group (previously classified as nonparticipants) could be subdivided into two categories: the first category included those who reported first UB participation in a grade prior to their grade placement at time of sampling. If their follow-up questionnaire responses were correct, then they were probably not defined in the original sampling frame (e.g., were not active UB participants at time of sampling and were not in the comparison population due to prior UB participation). Other explanations are possible. The second category contained those who reported first UB participation in their grade at time of sampling or subsequently. If their follow-up questionnaire responses were correct, then this group likely represents those who were nonparticipants at time of sampling but who joined UB after the time at which the sample was defined. For technical reasons, it was decided to exclude these groups also from the first follow-up analyses.[5]

---

5 These groups were not large enough to be included as a separate analysis category, and these cases represented certain sampling weight anomalies that might have distorted results had they been included with the original treatment group.

While those who were sampled as participants, but who no longer report UB participation, represent an interesting group, it was assumed on the basis of first follow-up frequencies (N = 37) that such individuals represented such a small segment of the population and such a distorted one that their exclusion from second follow-up analysis was justified. The nonparticipant-to-participant crossovers were expected to constitute a larger group (N = 193 in the first follow-up), and one that was legitimate for consideration, particularly in light of the participation pattern categories considered.[6]

Consideration of crossovers did represent technical problems, however. Since the originally defined comparison population was appreciably larger than the UB population and since selection rates from that population were appreciably smaller, the selection probabilities for comparison sample students were considerably smaller than for the corresponding Upward Bound students. Consequently, the 210 comparison students who subsequently joined the UB program had much larger sample weights (reciprocal of selection probabilities) than the original UB students. Incorporating these Comparison-to-UB crossovers directly into the treatment group estimates could have significantly reduced the stability of the resulting combined statistics.

RTI had recently experimented with several weight adjustment techniques that seemed particularly suited to the problem outlined above. This class of adjustment algorithms can be characterized as weight smoothing techniques that reapportion the total sampling weight sum between the "large" and "not-large" weight groups so as to control the resulting bias in combination statistics while attempting to minimize the inflationary effect of unequal weighting. The particular algorithm chosen for this study was relatively simple and could be implemented more efficiently than some of the more complex, more expensive, and only marginally more effective algorithms.

II. SPECIFIC PROCEDURES USED IN PARTIAL BALANCING

The procedures used in partial balancing involved an attempt at standardizing the characteristics of the comparison group of nonparticipants to the UB population characteristics over of the three initial grade cohorts 10-12.[7] Three separate partial balancings were implemented; one for the total group, one for the PSE-entering subgroup, and one for the subset of the PSE-entering group with available data for the structural equation solution reported in Chapter 5.[8] To better understand the balancing approach used in the current study, particularly its partial nature, the full balancing procedure that was used in previous studies will be described first. The full procedure is

---

6 Crossovers could define particular participation patterns that were theoretically impossible for those originally selected as participants (see Appendixes E and H).

7 It should be noted that the balancing procedure used in this study (as well as that used in previous studies) only attempted adjustment for two groups--former participants and nonparticipants. Adjustment within the former participant group on the basis of participation pattern was not attempted.

8 The two former balancing procedures were carried out in anticipation of a set of analyses that was not completely implemented.

described for balancing within grade cohort; however, an extension to balancing over grade cohort is quite straightforward.

A. The Full Balancing Procedure Within Grade Cohort[9]

The procedure used in previous studies involved balancing of the former participant and nonparticipant populations to be basically equivalent in respect to characteristics. This was accomplished by defining balancing groups of individuals with similar characteristics and by then adjusting the two groups (former participants and nonparticipants), so that the adjusted weighted proportional representation of the characteristics within each group was the same.[10]

For each of grades 10, 11, and 12, 16 balancing groups were formed both for UB participants and for the comparison group. The 16 balancing groups were based on two variables associated with UB participation qualifications, namely poverty status and academic risk, and on two additional variables, ethnicity and sex. The 16 balancing groups were formed as a partial crossing of these four variables as indicated in Table G.1.[11]

Within each of the three grades separately and for each of the 16 balancing categories, the total estimated relative population size was obtained for UB students and for comparison students. Using these numbers, a balancing weight component was computed for each category. Each comparison student's final sample weight was then multiplied by each balancing weight component appropriate to his categorization to obtain his two final balanced weights. Balanced or standardized rates, proportions, etc. could then be estimated for the comparison students by using the appropriate final balanced weights rather than the original sample weights. Comparing UB participant statistics with balanced comparison statistics would then be free of the effects of differential distributions among the balancing groups.

Only the within-grade balancing procedure is considered, since it is somewhat simpler to describe than the over-grade balancing, and since the latter balancing is completely analogous. The computations for within-grade balancing are given below:

---

[9] In general, subscripts and superscripts are used consistently in this subsection as follows:

- j, indicates an individual;
- $k \leftarrow 1, 2, \ldots, 16$, indicates a balancing category;
- $g \leftarrow 10, 11, 12$, indicates initial grade cohort;
- u, indicates former Upward Bound participant;
- c, indicates nonparticipant; and
- $'$ is used to indicate a balanced value.

[10] Original sampling weights were not used in the partial balancing implemented in the current study, thus, the balancing was restricted to the sample (or subsets of it), rather than to the population.

[11] This is effectively a balancing on the joint distribution of the variables considered. Such a balancing approach accounts for any interrelatedness of the variables used and thus ensures that the groups are also balanced on the marginal distributions of the variables.

Table G.1

DESCRIPTION OF THE SIXTEEN CATEGORIES USED IN BALANCING WITHIN GRADE

Ethnicity and Sex	Poverty Status and Academic Risk			
	Poverty Risk	Poverty "Not Risk"	"Not Poverty" Risk	"Not Poverty" "Not Risk"
Black, male	1	2	3	4
Black, "not male"	5	6	7	8
White, all sexes	9	10	11	12
Other, all sexes	13	14	15	16

NOTE: Table entries indicate the 16 balancing group numbers as a function of the marginal classifiers. The categories "not male," "not risk," and "not poverty," include cases which could not be classified due to indeterminate data as well as cases that could be classified as female, not risk, and not poverty, respectively.

For a given grade g, g ← 10, 11, 12, define the estimated number of UB students in grade g and in balancing category k, k ← 1, 2, ..., 16, as $Y_{ug}(k)$, which was determined as follows:

$$Y_{ug}(k) = \sum_{j=1}^{N} W_j U_j G_j K_j E_j ;$$

where

N is the total number of individuals (N=6050),

W_j is the eligible-adjusted weight for individual j,

U_j is a UB status indicator, with value 1 if individual j indicated UB participation and 0 otherwise,

G_j is a grade status indicator, with value 1 if initial grade cohort of individual j was g and 0 otherwise,

K_j is a group status indicator, with value 1 if the balancing group of individual j was k and 0 otherwise,

E_j is an eligibility indicator, with value 1 if individual j was eligible for the survey and 0 otherwise.

Also define, for grade g, the estimated number of nonparticipants in grade g and balancing category k as $Y_{cg}(k)$, which was determined as follows:

$$Y_{cg}(k) = \sum_{j=1}^{N} W_j \; C_j \; G_j \; K_j \; E_j \; ,$$

where C_j is $1-U_j$ and where N, W_j, G_j, K_j, and E_j are defined above.

Now $\pi_{ug}(k)$ the proportion of the grade g UB population in balancing class k, is defined as

$$\pi_{ug}(k) = \frac{Y_{ug}(k)}{\sum_{k=1}^{16} Y_{ug}(k)} \; .$$

An analogous proportion, $\pi_{cg}(k)$ was also determined for the comparison group as

$$\pi_{cg}(k) = \frac{Y_{cg}(k)}{\sum_{k=1}^{16} Y_{cg}(k)} \; .$$

Define the balancing weight correction component for category k as $\omega_g(k)$, where

$$\omega_g(k) = \frac{\pi_{ug}(k)}{\pi_{cg}(k)} \; .$$

The within-grade balanced weight for an individual, $\beta_g(j)$, was then determined within grade g and balancing category k as follows:[12]

$$\beta_g(j) = \begin{cases} W_j \; E_j, \text{ if } C_j = 0, \\ W_j \; E_j \; \omega_g(k), \text{ if } C_j = 1. \end{cases}$$

B. Partial Balancing Procedures

Since balancing category group sizes were restricted in the current study, due to the fact that more variables and partitionings were used, the full balancing approach could not be implemented (i.e., the number of cases within a particular balancing category and grade was in some cases reduced either to zero or to a size insufficient to provide stable estimates of the

---

[12] Note that the weight of UB participants (C_j=0) is not affected by the balancing. The use of the value E_j in this process is somewhat redundant, since it only serves to set the weight of ineligibles to 0. This had already been accomplished at the time balancing operations were implemented, with any appropriate adjustments (see Section VI of this appendix).

balancing weight correction component). Thus, a partial balancing was implemented. In general, the partial balancing approach differs from full balancing in that the full joint distribution of the balancing variables is not considered. Rather, the procedure involves only a consideration of either the marginal distribution of the separate balancing variables or marginal joint distributions of subsets of the variables. Thus, more than one set of proportions and more than one weight correction component is considered.

As such, each individual is categorized two or more times, such that categorizations are mutually exclusive and exhaustive (i.e., each categorization is defined by a different unique subset of the total set of balancing variables). Subsequently, adjustments were made by considering simultaneously the adjustments necessary for balancing within each of the separate categorizations. Unlike the full balancing procedure, the partial procedure does not guarantee equivalence of the marginal proportions, unless the separate categorizations are independent. The specific computational procedure within a particular grade is given below; considering the original sampling weight W_j.[13]

For each of M categorizations, define $Y^m_{ug}(k_m)$ and $Y^m_{cg}(k_m)$ as before, where $k_m \leftarrow 1, 2, \ldots, K_m$, and K_m is the number of distinct categories within the m^{th} categorization. Similarly, the proportion of UB participations in grade g and in balancing category k_m of categorization m is given by

$$\pi^m_{ug}(k_m) = \frac{Y^m_{ug}(k_m)}{\sum\limits_{k_m=1}^{K_m} Y^m_{ug}(k_m)} .$$

The corresponding proportion for nonparticipants, $\pi^m_{cg}(k_m)$, is given by an analogous equation in terms of $Y^m_{cg}(k_m)$.

The balancing weight correction component for category k_m, $\omega^m_g(k_m)$ is defined as

$$\omega_g{}^m(k_m) = \frac{\pi^m_{ug}(k_m)}{\pi^m_{cg}(k_m)} .$$

Such correction factors are computed for all K_m categories within each of the M categorizations considered.

The within-grade, partially balanced weight, $\beta'_g(j)$, for an individual who is classified in categories $k_1, k_2, \ldots, k_M$ may be determined as follows.

---

13 While within grade balancing was never implemented in the current study and original sampling weights were not used, the computational equations are presented in terms of these factors for comparability with the full balancing procedure equations.

$$\beta_g'(j) = \begin{cases} W_j\, E_j, \text{ if } C_j = 0, \\ W_j\, E_j \prod_{m=1}^{M} \omega_g^m(k_m), \text{ if } C_j = 1. \end{cases}$$

It can be observed from this expression and the previous discussion that if the M categorizations are independent, then the full and partial balancing approaches should lead to identical results, since the product of the corrections for the separate categorizations is the expected correction for the full joint categorization in the case of independence. That is, given independence,

$$\begin{aligned} \prod_{m=1}^{M} \omega_g^m(k_m) &= \frac{\pi_{ug}^1(k_1)}{\pi_{cg}^1(k_1)} \cdot \frac{\pi_{ug}^2(k_2)}{\pi_{cg}^2(k_2)} \cdots \frac{\pi_{ug}^M(k_M)}{\pi_{cg}^M(k_M)} . \\ &= \frac{\pi_{ug}(k)}{\pi_{cg}(k)} , \\ &= \omega_g(k). \end{aligned}$$

Under the more likely case of nonindependence, results of the two approaches are different and the partial approach does not guarantee equality of marginal proportions for the variables considered, or for any marginal joint distribution. Better solutions could probably be obtained through an iterative approach to partial balancing, but such an approach was not considered in this study.

C. Specific Modifications of the Partial Balancing Procedures

The major purpose of the balancing procedure used in the current study was to produce artificial groups of participants and nonparticipants that were: (1) similar in terms of the characteristics of grade cohort; poverty status; risk status; sex; race; and, in some cases, over-age-in-grade status, and (2) similar in size. To accomplish this in a way that could be accomodated by available software within the exploratory focus of the analyses involved, it was decided jointly by RTI and OED to examine the sample-specific data, without concern for sampling probabilities. While this precluded generalizability of related analytic results beyond the sample, it was a convenient approach within the accepted analytic focus, and it reduced problems with available software (see Appendix I).

Within the adopted approach, weights for both former participants and nonparticipants were set to unity. Then, within each of three distinct sets of sample members, a modified partial balancing approach was implemented. First correction factors were calculated, analogously to the procedure described in the previous subsection but using the unit weights. Next, a correction for differences in group sizes, ζ, was computed as

$$\zeta = \frac{\sum_{j=1}^{N} W_j\, U_j}{\sum_{j=1}^{N} W_j\, C_j} ,$$

Table G.3

GROUP PROPORTIONS AND ADJUSTMENT FACTORS FOR THE TWO CATEGORIZATIONS USED IN PARTIAL BALANCING FOR PSE-ENTERING SUBGROUP

First Categorization						Second Categorization					
Categories						Categories					
Poverty Status	Academic Risk Status	Race	Proportion of Participants	Proportion of Non-Participants	Adjustment Factor	Sex	Grade Cohort	Age-Within-Grade Status	Proportion of Participants	Proportion of Non-Participants	Adjustment Factor
Poor	Risk	Black	.219	.121	1.813	Male	10	Not Over	.044	.146	0.300
Poor	Risk	White	.030	.045	0.679	Male	10	Over	.005	.012	0.444
Poor	Risk	Other	.043	.028	1.542	Male	11	Not Over	.141	.133	1.059
Poor	Not Risk	Black	.294	.113	2.595	Male	11	Over	.024	.017	1.399
Poor	Not Risk	White	.075	.067	1.122	Male	12	Not Over	.177	.153	1.159
Poor	Not Risk	Other	.077	.030	2.546	Male	12	Over	.027	.017	1.604
Not Poor	Risk	Black	.068	.059	1.154	Female	10	Not Over	.064	.113	0.567
Not Poor	Risk	White	.012	.104	0.112	Female	10	Over	.005	.012	0.444
Not Poor	Risk	Other	.018	.024	0.764	Female	11	Not Over	.203	.166	1.225
Not Poor	Not Risk	Black	.103	.067	1.540	Female	11	Over	.012	.012	1.035
Not Poor	Not Risk	White	.036	.308	0.118	Female	12	Not Over	.274	.213	1.287
Not Poor	Not Risk	Other	.024	.034	0.700	Female	12	Over	.023	.006	3.461

NOTE: These values were determined using unit weights for the 760 nonparticipants and 1,713 former participants with determinate values of the classifying variables; therefore, cell sizes can be determined from the proportions given. The group size correction factor is seen to be 1,713/760 = 2.254

Table G.4

GROUP PROPORTIONS AND ADJUSTMENT FACTORS FOR THE TWO CATEGORIZATIONS USED IN PARTIAL BALANCING FOR PSE-ENTERING SUBSET WITH FULL DATA FOR PATH ANALYSIS

Categorization	Category	Proportion of Former Participants	Proportion of Nonparticipants	Adjustment Factor
Race	Black	.677	.301	2.253
	White	.151	.593	0.254
	Other	.172	.106	1.620
Academic Risk Status	Not Risk	.640	.686	0.933
	Risk	.360	.314	1.147
Grade Cohort	10	.112	.303	0.371
	11	.401	.327	1.227
	12	.487	.370	1.316
Poverty Status	Not Poor	.274	.665	0.412
	Poor	.726	.335	2.167
Sex	Male	.435	.497	0.875
	Female	.565	.503	1.124

NOTE: These values were determined using unit weights for the 376 nonparticipants and 917 former participants with determinate values of the classifying variables; therefore, cell sizes can be determined from the proportion given. The group size correction is seen to be 917/376 = 2.439.

of weight adjustment. Therefore, no corrections for partial balancing were considered in the analytic applications of these adjusted weights. Consequently, the results are suspect even under the adopted exploratory analysis focus.

III. WEIGHT ADJUSTMENTS FOR SUBSAMPLING AND NONRESPONSE[17]

The weight adjustment techniques described in this section are those that were implemented to compensate for the fact that data were not collected for all 6,050 individuals in the original sample. The weight adjustments discussed here involve: (a) corrections for telephone subsampling and (b) corrections for indeterminate response.[18] The various weight adjustment techniques (including balancing) were applied sequentially, and the order in which they were applied was important in determining an individual's final weight. The order of weight adjustments is presented in Section V of this appendix.

Clearly, some adjustment in weights was required due to subsampling of mail nonrespondents, since the telephone responses of those selected were used to represent not only themselves but also those mail nonrespondents who were not selected. Since the sampling within stratum was random, the subsampling adjustment procedure involved multiplying selected nonrespondents' weights by the inverse of the appropriate within-stratum sampling rate (see Appendix C, Table C.4), and setting the weights of nonselected nonrespondents to zero.[19]

The correction for instrument and item nonresponse (or indeterminacy) is not as immediately obvious but is equally important. It is reasonable to expect that the aggregate values of responses of individuals for whom data were not collected would have been different from the aggregate values of those for whom data were collected (i.e., nonresponse is not considered to be

---

17 In the following presentation, subscripts and superscripts are used consistently as follows:

- j, indicates an individual
- $k \leftarrow 1, 2, \ldots, 24$, indicates a weight adjustment category (class);
- $h \leftarrow 1, 2, \ldots, 27$, indicates project stratum;
- $i \leftarrow 1, 2$, indicates project within stratum;
- $g \leftarrow 10, 11, 12$, indicates initial grade cohort;
- u, indicates former Upward Bound participant (typical participation);
- v, indicates former Upward Bound participant (atypical participation);
- c, indicates nonparticipant;
- ℓ, indicates the mail nonrespondent sampling stratum;
- $W_\cdot(\cdot)$, indicates a weight prior to adjustment; and
- $W'_\cdot(\cdot)$, indicates a weight following adjustment.

18 An additional form of weight correction was employed involving the exclusion from analyses of individuals determined survey-ineligible, specifically setting the weights of such individuals to 0. This correction is straightforward and is not described in this section; however, it is referenced in Section V of this appendix.

19 The expected value of the adjusted weight totals over all possible subsamples is the unadjusted weight total, but, for a particular subsample, equivalence of adjusted and unadjusted weight totals is neither guaranteed nor expected.

a random process, and, therefore, is not independent of the study variable set). For example, the following assumptions are not unreasonable. The more mobile an individual, the less likely he would be located for data collection purposes and the less likely he would be expected to have obtained higher levels of educational progress. Likewise, the less satisfied an individual is with his educational progress, the less likely he would be willing to report it.

For cases in which nonrespondents and respondents are expected to differ, spurious results could obtain if one were simply to ignore nonrespondents or to assign their weights equally to all respondents. The procedure employed involved proportionally assigning the weights of nonrespondents to those respondents who are most like them. For the purpose of this study, the variables involved in determining the similarity of individuals include UB participation pattern (typical, atypical, or nonparticipant), grade, project sampling stratum, weight adjustment class, and UB project association.[20] The weight adjustment classes are related to risk, poverty status, ethnicity, and sex, as indicated in Table G.5.[21]

A. Specific Procedures Used in Adjusting for Subsampling

This adjustment was performed exclusively for the subset of individuals who were mail nonrespondents or undeliverables at the time of telephone sampling.[22] Since sampling rates were different among the ten nonempty strata used for the subsampling (see Appendix C, Table C.4), adjustments were made within those strata.

Within each telephone subsampling stratum, ℓ, the stratum sampling rate, R_ℓ was determined as

$$R_\ell = \frac{A_\ell}{A_\ell + B_\ell}$$

where

A_ℓ = the number of selected mail nonrespondents in stratum ℓ,

and

B_ℓ = the number of nonselected mail nonrespondents in stratum ℓ.

---

20 Nonparticipants were associated with a particular UB project through the association of their high school with the project. The unadjusted and adjusted weight totals will be identical, within rounding error, under this adjustment procedure.

21 The unadjusted and adjusted weight totals will be identical, within rounding error, under this adjustment procedure.

22 Certain individuals were excluded from the sampling frame and a weighting class adjustment was performed for this exclusion category prior to applying the subsampling adjustment.

Table G.5

DESCRIPTION OF THE TWENTY FOUR CLASSES USED IN WEIGHT ADJUSTMENT

Ethnicity	Sex	Poverty		"Not Poverty"	
		Risk	"Not Risk"	Risk	"Not Risk"
Black	Male	1	2	3	4
	"Not Male"	5	6	7	8
White	Male	9	10	11	12
	"Not Male"	13	14	15	16
Other	Male	17	18	19	20
	"Not Male"	21	22	23	24

NOTE: Table entries indicate the 24 weight adjustment class numbers as a function of marginal classifiers. The categories "not male", "not risk," and "not poverty" include cases which could not be classified due to indeterminate data as well as cases that could be classified as female, not risk, and not poverty, respectively.

The weight adjustment factor for stratum ℓ, ω_ℓ, was computed as

$$\omega_\ell = \frac{1}{R_\ell} \quad ,$$

and the telephone sample selection adjusted weight for individual j in stratum ℓ, $W'_{\ell j}$, was defined as

$$W'_{\ell j} = \omega_\ell W_{\ell j} S_{\ell j} \quad \text{where}$$

$W_{\ell j}$ = the unadjusted[23] weight for individual j in stratum ℓ,

and

$$S_{\ell j} = \begin{cases} 1 & \text{if individual j was a selected mail nonrespondent in stratum } \ell, \\ 0 & \text{if individual j was a nonselected mail nonrespondent in stratum } \ell. \end{cases}$$

The strata employed, number of subjects within each stratum, number sampled, and sampling rates (R_ℓ) are presented in Appendix C (Table C.4).

23 This was not the raw sampling weight for the individual, but reflected other weight adjustments (e.g., study eligibility) made to raw weights prior to the specific adjustment for telephone sample selection.

B. Specific Procedures Used in Adjusting for Nonresponse or Indeterminacy

In cases of nonresponse or other forms of indeterminate data, weight adjustments were made. In general, such adjustments involved apportioning the sampling weights of nonrespondents proportionally among the respondents who were in a category most like them (e.g., to those responding individuals from the same project or set of schools and of the same sex, race, academic risk status, poverty status, grade level, and UB participation group). The equations given are general and apply to any given adjustment of this type. Most adjustments were performed separately within the typical participation pattern, atypical participation pattern, and nonparticipant groups; however, some adjustments (i.e., those prior to adjustment for indeterminacy of participation pattern) were performed using only two groups, former participants and nonparticipants. For purposes of parsimony the equations provided are for the typical participation pattern group only (completely analogous equations would apply for both remaining groups or for the combined former participant group).

Define the unadjusted[24] weight totals of typical UB participant respondents and nonrespondents in grade cohort g and weighting class k of project i within stratum h as $T^r_{ugk}(hi)$ and $T^n_{ugk}(hi)$, respectively. These totals are given by

$$T^r_{ugk}(hi) = \sum_{j=1}^{n_{ugk}(hi)} R_{ugk}(hij)\, W_{ugk}(hij),$$

and

$$T^n_{ugk}(hi) = \sum_{j=1}^{n_{ugk}(hi)} [1-R_{ugk}(hij)]W_{ugk}(hij) ,$$

where

$W_{ugk}(hij)$ = the unadjusted weight of former typical UB participant j in project hi grade cohort g and weighting class k,

$$R_{ugk}(hij) = \begin{cases} 1 & \text{if former typical participant j in project hi grade cohort g and weighting class k was a respondent,} \\ 0 & \text{otherwise,} \end{cases}$$

and

$n_{ugk}(hi)$ = the number of typical former participants in project hi grade g and weighting class k.

The weight of typical former participant nonrespondents was apportioned to typical former participant respondents--within project hi, grade g, and

[24] Unadjusted in respect to the adjustment under consideration.

weighting class k--proportional to the respondents' unadjusted weights to give the adjusted weight, $W'_{ugk}(hij)$, for individual j. The adjusted weight was therefore defined as

$$W'_{ugk}(hij) = R_{ugk}(hij)\ W_{ugk}(hij) \left[1 + \frac{T^{n}_{ugk}(hi)}{T^{r}_{ugk}(hi)} \right] .$$

Note that the adjustment was not defined for the case in which there are nonrespondents but no respondents in project hi, grade g, and weighting class k; however, when the respondent group exists,

$$\sum_{j=1}^{n_{ugk}(hi)} W'_{ugk}(hij) = T^{r}_{ugk}(hi) + T^{n}_{ugk}(hi) = \sum_{j=1}^{n_{ugk}(hi)} W_{ugk}(hij).$$

For the case in which the procedure is undefined, it was necessary to change the definition of the group "most like" the nonrespondents. This process of redefinition was approached systematically and is set forth in the following subsection.

C. Search Procedure To Identify Similar Groups Containing Respondents

For those cases in which a particular group defined by UB participation category, grade cohort, stratum, risk group, poverty status, ethnicity, sex, and project within stratum contained nonrespondents but no respondents, a search process was initiated to locate a "similar" group containing respondents over which the nonrespondents' weights could be distributed. The first step in the search process was to determine whether or not respondents were present in the same UB participation category, grade cohort, stratum, and weighting class in the other project within the stratum.[25] If so, weights for nonrespondents were apportioned within that group in the manner described previously. If not, the search continued to see if respondents were present in the original project in the weighting class defined by the opposite sex.

This process continued until a sufficiently matching group was found or until it was determined that no respondents were present within the larger group defined by UB participation pattern, grade cohort, and stratum. Where this situation obtained, the weight total for nonrespondents was apportioned across respondents in the same major class, where major class is defined by the variables above with the exception of stratum. Only if no respondents were found in one of these groups were no weighting adjustments made and the weights of such nonrespondents lost.[26]

The ordered importance of weighting class definitions employed in this process was academic risk, poverty status, ethnicity, sex, and project. The

---

[25] Recall that two projects were selected within each of 27 strata.

[26] This situation was brought about through an anomaly in the weight adjustment software, which was discovered only after analyses were performed. It had been intended to apply the similarity algorithm again once stratum had been discarded; however, this feature of the weight adjustment program failed.

procedure of determining similarity within this ordering is fairly straightforward and is defined below.

Identify a particular subgroup within UB participation pattern-grade cohort-stratum class as follows:

$$G_{mnopqi},$$

where

m designates risk status $m = \begin{cases} 0 \text{ if non-risk,} \\ 1 \text{ if risk;} \end{cases}$

n designates poverty status $n = \begin{cases} 0 \text{ if non-poverty,} \\ 1 \text{ if poverty;} \end{cases}$

o designates white ethnicity $o = \begin{cases} 0 \text{ if non-white,} \\ 1 \text{ if white;} \end{cases}$

p designates black ethnicity $p = \begin{cases} 0 \text{ if non-black,} \\ 1 \text{ if black;} \end{cases}$

q designates sex $q = \begin{cases} q \text{ if male,} \\ 1 \text{ if female;} \end{cases}$

i designates project within stratum $i = \begin{cases} 0 \text{ if project 1,} \\ 1 \text{ if project 2.} \end{cases}$

Let this subgroup contain nonrespondents but no respondents.

Define a complement pattern as a six digit binary number, initialized to 000000, each digit of which corresponds to a subscript in the natural order. The search for a matching group containing respondents then proceeds as follows:

1. Add 1 to the complement pattern.

2. Modify each subscript based on the value of its associated complement pattern digit:

   0 → no modification,
   1 → complement.

3. If respondents are present in the newly-identified subgroup, distribute nonrespondent weights from the original group over respondents in the new group; if not, return to Step 1.

If no respondents were found within a particular class, the procedure is repeated using an analogous weighting class definition that ignores project and stratum altogether.

As an example of this procedure, consider that a subgroup characterized by: (1) non-risk, (2) non-poverty, (3) non-white, (4) non-black, (5) female, (6) project 2, has only nonrespondents. Following the notational scheme set forth above, this subgroup is designated as:

G_{000011}.

The complement pattern initialized to 000000 is incremented by 1 to yield 000001. The application of this complement pattern t^ the subscripts yields

G_{000010}

as the "most similar" group. This corresponds to a change in only project membership from project 2 to project 1. If no respondents are present in this group, the complement pattern is increased by 1 to yield 000010. This complement pattern, when applied to the original subscripts yields

G_{000001}

as the "next most similar" group. The most similar group is now defined as male respondents in the original project. Assuming no respondents exist within this group, the complement pattern is increased by 1 to yield 000011, and application to the original subscripts creates a most-similar respondent group

G_{000000}

(i.e., project 1, sex male).

This process continues until a group containing respondents is found and weight adjustment performed or until the complement pattern is incremented beyond 111111 at which point an informative message appears and an analogous procedure is implemented ignoring stratum and project. The only peculiarity to this procedures occurs when o and p are both equal to 1, corresponding to a null subgroup; however, the developed procedure automatically excluded this subpattern from consideration.

IV. ADJUSTMENTS FOR NONPARTICIPANT CROSSOVERS

Weight smoothing techniques were used when adding to the former participant group those who were originally sampled as nonparticipants, but who subsequently reported UB program participation (crossovers). As stated previously, the originally defined comparison population was appreciably larger than the UB population and selection rates from that population were appreciably smaller; thus, the selection probabilities for comparison sample students were considerably smaller than for the corresponding UB sample students. Incorporating these crossovers directly into the treatment group estimates would have significantly reduced the stability of the resulting combined statistics. While the weight smoothing technique employed probably resulted in some bias in the estimators, it has the effect of enhancing estimate stability (e.g., minimum variance estimators, in which the variability due to both sampling fluctuations and bias is reduced).

The motivation for the weight adjustment strategy may be seen from a simplified model for the form of combination statistics which presumes that the relative size of the large and not-large weight groups is known. That is, if π represents the true fraction of Upward Bound participants in the target population that were originally in the comparison universe, then a weighted sample mean combining both types of UB participants would have the form

$$\bar{x}_\pi = \pi \, \bar{x}_L + (1-\pi)\bar{x}_{NL}$$

where $\bar{x}_L$ and $\bar{x}_{NL}$ denote sample means related to a specific outcome measure for the large weight (crossovers) and not-large weight (original UB) students respectively.

While it is recognized that in most instances one would have to estimate π, using the weighted proportion of group L members in the combined sample, the assumption of known π leads to a manageable variable approximation for $\bar{x}_\pi$, which should suffice for current purposes; namely,

$$\mathrm{Var}(\bar{x}_\pi) \doteq \pi^2 \, V_L + (1-\pi)^2 \, V_{NL} + 2\pi(1-\pi) \, C_{L,NL} ,$$

where V_L and V_{NL} are the sampling variances of $\bar{x}_L$ and $\bar{x}_{NL}$, respectively, $C_{L,NL}$ is the sampling covariance between $\bar{x}_L$ and $\bar{x}_{NL}$ that can be induced by cluster sampling designs, and where both the L and NL samples are drawn from within a set of primary sampling units.

The effect of large weights will be manifested by a large contribution from the leading term in the variance expression given above. When the sample size n_L associated with the large weight group is considerably smaller than its proportionate share of the combined sample, one can expect serious variance inflation unless the group L outcome measures are correspondingly less variable.

To illustrate this point, consider a simplified sample design where comparison and UB students were selected by simple random sampling from national listings. For such a simplified sampling scheme, a mean PSE outcome measure for original nonparticipants with subsequent UB exposure would have approximate variance $V_L \doteq S_L/n_L$ where S_L represents population variation in outcomes among all nonparticipant list members with later UB participation. Similarly, $V_{NL} \doteq S_{NL}/n_{NL}$ would approximate the variance for the original UB sample mean. The cluster induced covariance $C_{L,NL}$ would be eliminated within this simplified sampling model.

Assuming further that the PSE outcome variations are roughly equivalent for the two groups, namely $S_L = S_{NL} = S$, then

$$\mathrm{Var}(\bar{x}_\pi) \doteq [\pi^2/\theta + (1-\pi)^2 \, / \, (1-\theta)]S/n_+$$

approximates the variances of the π weighted combination of PSE outcome means, with n_+ representing the combined sample size and θ the group L fraction of the combined sample, namely $\theta = n_L/n_+$. The quantity in brackets (representing the inflationary effect of unequal weighting) exceeds one, since θ, the sample allocation to the large weight group, is considerably smaller than π, the population allocation of UB participants to group L.

In the current illustration, where the average weight for original UB students is 7.7 compared with an average of 936.0 for original CS students, the 210 comparison students with subsequent UB experience would inflate to 196,560 nationally. Thus, their responses could overwhelm responses for the 2,678 first follow-up original UB respondents with a national expansion of 20,620. These expansions lead to an estimated population fraction $\pi = .901$ for the large weight crossovers whose first follow-up sample allocation is $\theta = 200/2,878 = .0695$. The associated variance inflation factor compared to an equally weighted total sample of 2,925 UB participants is 11.69 according to the simplified formulation, or roughly a 1,000 percent increase. Compared to estimates based on the sample of 2,678 original UB participants alone, combined statistics weighting in the 210 crossovers with no adjustment would have inflated the variance by a factor of 11.5.

While the previous discussion suggests that ignoring the crossovers in the first follow-up analysis was a prudent decision in terms of estimate stability, excluding them completely may be introducing more bias than is desirable. If μ_L and μ_{NL} denote the population means corresponding to the sample estimates $\bar{x}_L$ and $\bar{x}_{NL}$, the bias in $\bar{x}_{NL}$ as an estimate of the true mean $\mu = \pi\mu_L + (1-\pi)\mu_{NL}$ is

$$\mathrm{Bias}(\bar{x}_{NL}) = -\pi(\mu_L - \mu_{NL}) .$$

Since $\pi \doteq .9$ in this instance, most of the difference between μ_L and μ_{NL} is translated directly into bias. As an alternative to $\bar{x}_{NL}$ one may consider a general biased linear estimator

$$\bar{x}_\phi = \phi\,\bar{x}_L + (1-\phi)\bar{x}_{NL}$$

where

$$\mathrm{Bias}(\bar{x}_\phi) = (\phi - \pi)(\mu_L - \mu_{NL}).$$ [27]

Denoting the mean difference $(\mu_L - \mu_{NL})$ in the bias expression above by δ and treating the adjusted combining fraction ϕ as an arbitrary constant, the Mean-Square-Error (MSE) or squared bias plus variance for $\bar{x}_\phi$ has the form

$$\mathrm{MSE}(\bar{x}_\phi) = (\phi - \pi)^2\,\delta^2 + \phi^2\,V_L + (1 - \phi)^2\,V_{NL} + 2\,\phi(1 - \phi)\,C_{L,NL} .$$

The optimum value of ϕ which minimizes the mean square error expression above is

$$\phi_{opt} = (\delta^2\pi + V_{NL} - C_{L,NL}) \,/\, (\delta^2 + V_L + V_{NL} - 2\,C_{L,NL}) .$$

[27] In the special case used for the first follow-up analysis, ϕ took on a value of 0.

The optimum combining fraction derived above is defined in terms of unknown population means, variances, and covariances associated with a specific outcome measure; however, ϕ may by approximated by substituting sample estimates for these parameters. The large weight group has sufficient sample size in this instance to supply reasonable mean and varaniance estimates for the crossover domain. Along with the corresonding mean, variance, and covariance estimates for the original UB students, the transfer group estimates could be used to approximate the optimal weight deflation factors $\hat{\phi}/\hat{\pi}$. The original UB students would have their weights inflated by the compensating amount $(1-\hat{\phi}/\hat{\pi})/(1-\hat{\pi})$.

It should be noted, however, that the optimal adjustment factors are specific to a particular outcome variable. Thus, the adjustment factors would have to be computed for each analysis. Alternately, one could try to establish a global adjustment that performs adequately for all PSE outcomes. Both of these approaches would have been quite costly; therefore, other approaches to weight smoothing were considered.

The procedure selected was quite simple and easily implemented. The procedure involved: (1) determining the average weight for the total group of originally selected UB participants and crossovers; (2) assigning the average weight to all crossovers; and (3) apportioning the remaining weight total uniformly over the originally sampled participants. Although it had been planned originally to perform such smoothings within projects, such a procedure resulted in certain anomalies. Since weights for originally sampled participants within a particular project were identical (see Appendix C), except for minor differences introduced through previous adjustments, smoothed weights for original participants and crossovers would also be equivalent with project. Further, since the number of crossovers varied substantially over projects, smoothed weight sums within some projects (one with 10 crossovers) would have been markedly higher in some projects than in others (one with no crossovers). This could have inflated variances estimates between projects (both overall and within stratum) in a manner analogous to the variance inflation of student estimates considered previously. Thus, at the risk of introducing additional bias, the procedure was applied to the total groups, irrespective of project, in an attempt to further stabilize the estimates.

In defining this adjustment procedure, let

T_u = weight total for originally sampled UB participants,

T_t = weight total for crossovers,

n_u = number of originally sampled UB participants,

n_t = number of crossovers.

Then the smoothed weight, W_j', for individual j, may be expressed as a function of these values and the unadjusted weight as

$$
W_j' = \begin{cases} W_j, & \text{if individual is a noncrossover nonparticipant,} \\ \dfrac{T_u + T_t}{n_u + n_t}, & \text{if individual is a crossover,} \\ W_j + \dfrac{1}{n_u}\left[T_t - n_t \left\{\dfrac{T_u + T_t}{n_u + n_t}\right\}\right], & \text{if individual is an originally sampled UB participant.} \end{cases}
$$

It is easily seen that this procedure maintains the total weight sum for those involved in the smoothing.

VI. SUMMARY OF WEIGHT ADJUSTMENTS

The actual techniques employed in weight adjustments have been described in previous sections of this appendix; however, the order in which the various techniques were applied has not been addressed. Weight adjustment was a sequential process, and the nature of that sequence is important to an understanding of the basic character of the resultant weights used in the various analyses summarized in Chapter 3 of this report.[28]

There were nine basic steps in producing the various intermediate and analysis weights of this study. At each step, the weight resulting from the previous step was adjusted by a particular technique; thus, after step 1, the input to any of the adjustment techniques employed was a previously adjusted weight. Due to the basic differences in the adjustment techniques defined above, applying these three techniques in different orders will typically result in different weights. Moreover, some steps in the adjustment process involved removal of certain individuals from the data base (i.e., setting their weights to zero), and the stage at which such individuals were removed would affect the final weight distribution among those remaining.

The first step was adjustment of the raw sampling weights for second follow-up study ineligibility, involving 125 (e.g., 89 former UB participants and 36 nonparticipants). These cases included those who had been determined ineligible for the first follow-up survey[29] and additional individuals who had refused participation or had been reported deceased or incapacitated during the newsletter tracing or returns to the questionnaire mailing/tracing operation. Some of these cases were clearly not defined within the sampling frame of the base-year study, and the associated weights were set to 0. Other cases represented individuals who should have been available for followup but who had refused or could not be identified; sampling weights associated with these cases were therefore appropriately redistributed to other students in the sample (as described in Subsection III.B of this appendix). Deceased sample members were considered to be no longer in the population of interest, and their weights were set to zero.

---

[28] Since the weights used for analyses reported in Chapter 4 were not based on original sampling weights but rather on arbitrarily assigned unit weights, they will not be discussed in this section.

[29] See Appendix E of Burkheimer, G. J., _et al_., September 1977. (_op. cit._)

The second weight adjustment step involved a correction for the exclusion from the telephone sampling frame of 308 individuals who had been classified as "hard-core" undeliverables (see Appendix C). This group was comprised of 163 former UB participants and 145 nonparticipants, who should have been considered. Thus, their weights were distributed to similar individuals, as described in Subsection III.B above.

The third weight adjustment step consisted of adjustment for the subsampling of the remaining mail nonrespondents for telephone interview, using the procedure specified in Subsection III.A of this appendix. This adjustment step was confined to the subset of individuals who had not responded by mail at the time of subsampling and resulted in setting to zero the weights of 1,111 nonselected individuals (618 from the former UB participant group and 493 from the nonparticipant group) while increasing the weights of those selected by the appropriate stratum factor.

The fourth step in the process involved adjustment for "instrument nonresponse," which in this case was equivalent to a telephone refusal or a failure to contact by telephone. This adjustment used the procedures defined in Section III.B of this appendix and involved an apportioning of weights of 262 instrument nonrespondents (186 from the former participant group and 76 from the nonparticipant group) to individuals most like them.

The fifth step in the adjustment process involved the correction for 58 additional "ineligible" individuals (50 former participants and 8 nonparticipants) from the data base (i.e., setting their weights to zero). The specific individuals removed were: (a) those who were reported as deceased by telephone, and (b) those with UB participation classification conflicts or inappropriate classification at time of original sampling, as determined from questionnaire responses. Since those selected for telephone interview were to represent those not so selected and since those for whom information was collected by telephone represented those for whom telephone data were not collected, the deceased sample members were removed at this step (after adjustments for subsampling and noncontact) rather than at the earlier step in which other deceased sample members (reported by mail return) were removed.

The removal of cases with conflicting UB participation classification had been anticipated, based on first follow-up findings. The conflicts involved individuals who were originally selected into the participant group but who now indicated they never participated in UB. This group consisted of individuals who were either inappropriately defined in the sampling frame or who did not report their UB participation for some reason (forgot about it or did not wish to report it). Various forms of bias could have entered the results by inclusion of these cases (the type of bias suspected varies with the reason for the conflicting classification), and exclusion of the cases seemed justified.

The sixth step in adjustment was the weight smoothing technique described in Section IV of this appendix. While no cases were lost in this adjustment, 171 individuals[30] who had originally been sampled as nonparticipants were moved to the participant group as "crossovers." The weight total for these

[30] Of the 210 crossovers identified in this study, 39 had been excluded through previous weight adjustment.

171 cases was almost 6 times that for the 2,604 remaining originally sampled UB participants.

The seventh step in the weight adjustment process involved adjustments for nonresponse or indeterminacy (see Appendix D) of basic classifier variables for the analysis. These included extent of prior UB participation, initial grade cohort, sex, race, poverty status, academic risk status, and over-age-in-grade status. These adjustments involved the reapportioning of weights for a total of 831 individuals (181 nonparticipants and 650 former participants, including crossovers) with data indeterminacies on one or more of these variables by the procedure defined in Subsection III.B of this appendix.

The eighth step in weight adjustments involved corrections for the subsetting of the sample required in certain analyses (e.g., analyses that considered only a particular group of individuals, such as the PSE-entering group). There were several such adjustments, one for each unique subset defined by analysis, and these adjustments involved no more than setting to zero the weights of those individuals that were not included in the subset of interest.

The final step involved weight adjustments for indeterminacy of any additional variables (e.g., criterion variable and additional classifiers, if applicable) specific to each of the analyses involved. Thus, a separate adjustment was made for each analysis performed. The specific number of cases and weight totals for the separate analyses are provided in Appendix H.

Table G.6 presents a summary of the results of the first eight weight adjustment steps,[31] including summary statistics of the weights within the participant, crossover, and nonparticipant groups. This presentation clearly shows the shrinkage of the effective sample size during the several adjustment steps. The table also demonstrates the typically greater disparity between individual weights as more and more adjustments were made; statistical compensation for such disparities typically resulted in larger standard errors.

---

[31] For ease of presentation, only two of the several adjustments in Step 8 are provided.

Table G.6

SUMMARY OF WEIGHT ADJUSTMENTS

Step	Adjustment	Original UB Participants				Crossovers				Nonparticipants			
		Number Nonzero Weights	Weight Sum	Minimum Weight	Maximum Weight	Number Nonzero Weights	Weight Sum	Minimum Weight	Maximum Weight	Number Nonzero Weights	Weight Sum	Minimum Weight	Maximum Weight
1	Study Eligibility	3,621	20,657	2.59	10.31	210	120,832	18.32	4,118.34	2,094	1,277,808	18.04	7,394.32
2	Telephone Sample Frame Inclusion	3,458	20,657	2.59	22.62	207	124,691	18.32	4,118.34	1,952	1,273,948	18.04	7,394.32
3	Telephone Sampling	2,840	20,757	2.59	41.50	177	123,012	31.98	6,315.00	1,489	1,292,827	18.32	11,168.60
4	Telephone Nonresponse	2,654	20,756	2.59	46.40	172	123,345	31.98	6,315.00	1,418	1,292,493	18.32	11,168.60
5	Empirically Determined Study Eligibility	2,604	20,426	2.59	46.40	171	121,799	31.98	6,315.00	1,411	1,286,334	18.32	11,168.60
6	Weight Smoothing	2,604	133,460	46.00	89.80	171	8,764	51.25	51.25	1,411	1,286,334	18.32	11,168.60
7	Indeterminacy of UB Participation Pattern	2,500	133,241	46.00	175.80	171	8,983	51.25	109.70	1,386	1,286,334	18.32	11,168.60
7	Indeterminacy of Other Major Classifiers	2,015	136,139	46.00	994.80	110	7,849	51.25	198.20	1,230	1,284,571	18.32	12,425.90
8	Subsets of High School Graduates	1,969	133,067	46.00	994.80	107	7,682	51.25	198.20	1,168	1,220,330	18.32	12,425.90
8	Subset of PSE Entrants	1,641	110,265	46.00	994.80	70	5,262	51.25	198.20	746	760,281	19.33	12,425.90

Appendix H

Procedures and Results of Comparative/Descriptive Analyses

Appendix H

Procedures and Results of Comparative/Descriptive Analyses

I. GENERAL

This appendix particularizes the comparative/descriptive analyses leading to the findings reported in Chapter 3 of this report. Specifically emphasized are: (a) the general approach to analysis, (b) constraints and limitations of the analyses, and (c) tabular presentations of analysis results. Topics that are closely related to these analyses are detailed elsewhere: weighting adjustments, in Appendix G; and data reliability and analysis variable specifications, in Appendix E.

The comparative/descriptive approach is descriptive in the sense that estimates were produced, for both UB participants and nonparticipants, of group characteristics (e.g., pertinent experiences, characteristics of PSE institutions attended), and comparative in that group or subgroup differences on outcome or criterion variables were estimated as a function of the pattern of UB participation; in such comparisons, nonparticipants may be classified as individuals with a pattern of no participation in UB.

Each of the analysis questions of this study was addressed by these analyses. While the comparative/descriptive analyses were similar in form to the analyses of participant and nonparticipant differences performed in the base-year[1] and First Follow-up[2] studies, the Second Follow-up analyses differed considerably from previous ones in terms of the specific techniques used.

Using the student as the unit of analysis, aggregate differences in the outcome variable measures were examined within a general linear model as a function of pattern of UB participation, with statistical control (covariance adjustment) of a set of concomitant variables (specifically, student background factors).[3] Additional control was maintained in many analyses through performing analyses within specific subsets of individuals partitioned on the basis of other control variables (e.g., grade cohort, type of institution attended). The results of these analyses reflect the overall relationship

---

1 Burkheimer, G. J., Levinsohn, J. R., Koo, H. P., and French, A. M. Evaluation Study of the Upward Bound Program (Volume IV of A Study of the National Upward Bound and Talent Search Programs. Four Volumes). Research Triangle Park, N.C.: Center for Educational Research and Evaluation, Research Triangle Institute, April 1976. (ED 131-152)

2 Burkheimer, G. J., French, A. M., Levinsohn, J. R., and Riccobono, J. A. Evaluation Study of the Upward Bound Program: A First Follow-up. Research Triangle Park, N.C.: Center for Educational Research and Evaluation, Research Triangle Institute, September 1977.

3 The comparison population was defined as all nonparticipants enrolled in the tenth, eleventh, or twelfth grade of high schools served by a UB project in September 1973. There was no reason to believe, however, that the distributions of this population by grade or by relevant background characteristics would be exactly the same as those of the UB student population, and, as shown in Table H.1 of this appendix, the groups differed markedly on some such factors.

between participation in the UB program and student outcomes, controlling for differences in input to the program.

In prior comparative/descriptive analyses, statistical control for differences between participants and nonparticipants was obtained through a form of balancing.[4] The degree of additional statistical control required by the current study (e.g., control within each of the patterns of UB participation, and control within types of PSE institutions), could not be accommodated by the previously-used technique. Thus, the more common approach to statistical adjustment (covariance adjustment for main effects through a general linear model approach) was employed. Although more factors could be adjusted by this technique, it should be recognized that none of the interactions of those factors were considered, as had been accomplished by the previous technique.

In general, the comparative/descriptive analyses were restricted to the examination of single-degree-of-freedom contrasts of relevant means or proportions for a given outcome measure. Omnibus tests for main effects or interactions were not implemented; rather, meaningful differences in percentages or means between relevant subpopulations were considered. For each difference (single-degree-of-freedom contrast) an appropriate standard error was calculated, allowing for sampling effects, to provide confidence limits and thus a means of testing for statistical significance of the difference.

Some analyses were performed within initial grade cohort, since certain educational outcomes (e.g., educational progress) were restricted in the value they could assume by the sample members' educational starting point during the base-year study. Other analyses were performed for the total participation pattern groups, combined over grade cohorts, with appropriate adjustments. Additional partitioning of groups on other factors was undertaken to the extent possible, particularly in examining PSE outcomes when such factors (e.g., institutional level of offering) were considered to be logically related to the outcome under consideration. For most major analyses, the specific *a priori* contrasts constructed within a given partition of the groups were those reflecting major USOE interest (i.e., the contrasts between the typical and atypical participant groups and between the typical participant group and the nonparticipant group).

The group of individuals who were originally identified as nonparticipants but who subsequently reported participation in UB (subsequently called "crossovers") were considered for the first time in these analyses. Because of the large differences in analysis weights between former participants and nonparticipants (reflecting markedly different selection probabilities) and the fact that these individuals could no longer legitimately remain in the nonparticipant group, they had been excluded from first follow-up analyses. In the current study, a methodology for incorporating these individuals into the participant group was implemented (see Appendix G).

To provide unbiased estimates of population parameters, sampling weights (the inverse of the probability of selecting the individual) were used in

---

[4] After the comparison group had been adjusted to reflect the same distribution of sex, race, poverty status, academic risk status and grade cohort as the UB participant group, comparisons were made between the groups using software developed specifically for the UB evaluation study, which compensated for the adjustment technique. See Appendix E of Burkheimer, et al., 1977, (*op. cit.*).

analyses. The use of sampling weights overcomes possible bias in resulting statistical estimates which may be introduced because individuals were selected with unequal probability (due to oversampling of specific groups). Two conceptually different weight adjustment procedures were used to achieve more meaningful results (see Appendix G). The first procedure is a technique of weight smoothing that allowed consideration of the crossovers without unduly influencing the results by their disproportionately large sample weights. The second procedure involved the adjustment of sample weights for the sampling procedures used, for instrument nonresponse, and for item indeterminacy. The weight adjustment approach, which involved apportioning of sampling weights of nonselected and nonresponding individuals to those of selected and responding individuals, was chosen over several less conservative imputation approaches.

The basic approach used in analysis was multiple regression, a very general statistical technique whereby the relationship of one variable to a <u>set</u> of other variables may be assessed.[5] Generally, the technique allows the determination of that <u>linear combination</u> of those variables in the predictor set which best predicts (shows the greatest zero-order correlation with) the single criterion variable. In determining the best linear combination of the predictor variables, it is important to go beyond the simple zero-order relationship of each <u>individual</u> predictor to the criterion. Rather, the approach considers the relationship of a predictor variable to the <u>residual</u> portion of the criterion variable that remains when the prediction of <u>all</u> other variables in the predictor set has been removed (partialed).

Aside from the various theoretical and philosophical issues involving the validity of inferences from surveys, the basic problem of producing unbiased estimates of regression parameters and estimates of the associated standard errors has been a particularly thorny issue in dealing with results from stratified multistage sample designs such as the one used here. Most of the available statistical software packages treat the sample as independent random observations, ignoring the sample design. This approach is convenient, but it is inappropriate, since it does not account for unequal probabilities of selection. The application of sampling weights is possible through some software packages, allowing correct estimates of regression coefficients, but appropriate error variance estimates typically are not produced. In fact, it is not possible to obtain explicit expressions for variance estimates of complex estimators such as regression coefficients within complex survey sample designs; however, there are various approximation procedures available. While there is no single "best" way to approach the matter, RTI has developed an approach that is considered optimal or near optimal in most circumstances. The approach combines the use of the Horvitz-Thompson estimator[6] and Taylorized deviation[7] which have been incorporated into a statistical software routine.[8]

---

[5] See, for example, Draper, N. R. and Smith, H. <u>Applied Regression Analysis</u>. New York: Wiley, 1966.

[6] Horvitz, D. G. and D. G. Thompson. A Generalization of Sampling Without Replacement from a Finite Universe, <u>JASA</u>, 1952, 47, pp. 663-685.

[7] Folsom, Ralph E., Jr. <u>National Assessment Approach to Sampling Error Estimation</u>, Sampling Error Monograph, Project No. 25U-796-5, Research Triangle Park, N.C.: Research Triangle Institute, 1974.

[8] Holt, Mary M. <u>SURREGR: Standard Errors of Regression Coefficients from Sample Survey Data</u>, Research Triangle Park, N.C.: Research Triangle Institute, May 1977.

Although the Horvitz-Thompson estimator is not the "best" estimator in all cases, neither is any other estimator. Moreover, the Horvitz-Thompson estimator is widely used and has been found to be an admissible estimate. The Taylorized deviations approach is also not necessarily the "best" approximation procedure; but it does possess some practical and computational advantages over other available techniques (e.g., independent replications, pseudo-replications, or jackknife techniques). The Taylorized approach is, however, less known and is detailed in Subsection III.B of this appendix. While the procedure is an appropriate one, it imposes certain additional restrictions on analysis, particularly in regard to the degrees of freedom (and thus the number of parameters) available to the regression model and in regard to the use of variables in the model that were used in defining the original sampling strata. Specifically, models involving more than 25 regression degrees of freedom could not be analyzed by this procedure.

II. CONSTRAINTS AND LIMITATIONS OF ANALYSES

Care was taken throughout analysis to avoid techniques or comparisons that would introduce bias in results that was either favorable or unfavorable to the UB program. Also considerable effort was made to ensure that the data entering analysis were error free and consistent (see Appendix E). The original sample was designed to reduce many of the possible differences which could exist between the UB and comparison samples (and which, therefore, could lead to spurious group differences). Additionally, statistical control was undertaken in an attempt to reduce possible contamination of the results by existing differences between groups on background factors, initial grade distribution, and other factors. Weight adjustments for indeterminate response at the instrument and item level were employed in an effort to avoid distortion of the findings due to incomplete data collection. Standard errors were computed for all major differences so that probability statements concerning the populations of interest could be made and statistical significance of differences could be assessed.

Although it is quite likely that these efforts were at least partially successful, they did not ensure removal of all possible confounding factors, nor did they eliminate all constraints and limitations on the study findings. The analysis technique used is a relatively complex approach to general linear modeling and is subject to specific limitations. The technique assumes specific distributional properties of the variables analyzed, which were met only to varying degrees. Particularly, the assumptions of multinormal distributions and homoscedasticity of error are probably violated. Moreover, scale properties of some of the variables analyzed were not sufficiently strong to meet the assumptions of the analytic models (e.g., assumed ordinal scale properties were not always met). While none of the specific violations of model assumptions were considered to be critical (in fact, most represent generally accepted practices), the several violations, in combination, could lead to spurious results.

There are also several other potential sources for error in the statistics computed. Measurement error is an obvious source of such error and is addressed in Subsection II.A, below. Another source is sampling error, which is reflected in almost any estimate of a population characteristic from data collected from a sample of that population. This error is introduced by the fact that the computed estimate will vary from one specific sample to another. Sampling error is typically incorporated in statistical tests of reported differences, through one or another assumptions about the characteristics of the error under the sampling procedure used (see Subsection III.B). Sampling errors were not computed, however, for the point estimates computed within a specific participation pattern group. As a result, such point estimates should be considered only as approximations, varying from the true population values as a result of sampling error.

The potential for bias in the results also exists. Bias is typically introduced through a systematic misrepresentation of the population in the sample. While the full sample drawn in the base-year study would have led to unbiased estimates, this is not necessarily the case for subsets of the full sample. It is known that the weight smoothing technique applied to allow consideration of crossovers introduced bias into the results (see Appendix G). Also, when data collected include nonresponse and/or less than complete coverage of the sample, then the potential for bias exists when making population estimates from the subset of individuals who have provided data. It is quite likely that nonresponse or other factors leading to data unavailability (e.g., sample attrition since the base-year study) is not random and that bias has been introduced in some point estimates. There is little indication, however, that any nonrandom sample loss was systematically different for former participants and nonparticipants, thus differences between these two groups may show less bias than the estimates within each group, from which the difference is formed. If the bias is constant for both groups, then it effectively cancels itself when a difference is formed.

Two particularly difficult remaining confounds are associated with the relationships between pattern of UB participation and other variables used in the study. There are also specific constraints involving the statistical adjustments performed. Further, two basic problems inherent in questionnaire surveys remain: namely, questionable validity of reported information and the inability to definitively establish causation. These limitations on the study findings are considered below in some detail.

A. Validity of Data

Much of the data for this study is based on unverified self-reports or, in some cases, reports from secondary sources. In addition to known biases in student reports of academic achievement, invalid responses may be traced to a diversity of possible sources (e.g., deliberate attempts to mislead, lack of understanding or misinterpretation of questions, and simple errors in marking responses). Additional (or different levels of) errors are introduced when secondary source data are considered. Additionally, some error may be introduced in transcribing the data provided. No direct validation of collected data was attempted; validation studies, while theoretically very important, are typically not cost-feasible. Three efforts were attempted to assess data quality; two of these involved examinations of data reliability (a necessary but insufficient condition for validity). One examination involved the comparison of responses on duplicate questionaire returns; another involved checks for determinacy and internal consistency of questionnaire data. The final approach examined intercoder reliability in abstracting data from the PSE transcripts. These studies are detailed in Section VII of Appendix E.

Examinations of duplicate returns indicated a relatively high degree of agreement in reports of relatively objective information (e.g., high school completion, PSE entry) but considerably less agreement in reports of more subjective information (aspirations, problems experienced), reflecting in part the temporal instability of these latter measures. Also, this examination revealed a low reliability of reported amounts of financial aid received (which was anticipated on the basis of pretests of the questionnaire).

Data indeterminacy was considered to be within tolerance levels for most data elements; however, indeterminacy rates exceeded 15 percent for six variables considered. The items with greatest indeterminacy rates were those related to financial aid and educational expectations or aspirations. Some of the indeterminacy problems were "fixed" through logical imputations, introducing potential for another type of error into the results; however, in all but one instance (again educational aspirations) the amount of data that was imputed was less than 5 percent. Examination of data consistency was possible for some items of information. For the items considered, consistency was quite satisfactory in most instances; however, the consistency checks were only capable of detecting very large response discrepancies (see Appendix E). Greatest inconsistency existed in reported UB participation as compared to previously established participation. Although these inconsistencies may be somewhat artifactual,[9] they nevertheless signal the possibility of a considerable lack of reliability in the retrospective measure of program participation.

Although the data as it existed on student transcripts may be assumed quite reliable, the process of abstracting these data for the current study provided a source of considerable potential for error. The examination of intercoder reliability indicated that this potential had been realized in several cases. Although the estimated reliability of most major variables was satisfactory, particular problems were experienced with major course of study and with variables that were conditional on classification within a particular content area (e.g., GPAs within content area). In the latter case, reliability was so poor that certain planned analyses were not implemented.

The problem of data validity may have been compounded through the weight adjustment procedures, particularly if invalid data were collected from a student who was similar to a large group of nonrespondents. The real danger of the validity problem, however, is that of differential validity of responses within the two groups (UB and comparison). While a lack of reliability and validity serves to introduce error into the data, if this error is the same (regardless of direction or magnitude) within the UB and comparison groups, then group <u>differences</u> should show no directional bias. If, on the other hand, the data from one group are biased in a particular direction and those from the other group are unbiased (or biased in the opposite direction), then the bias will remain (or be magnified) in the group differences. None of the reliability or consistency checks were sufficient to address this issue, and, while there is no reason to suspect differential response biases, the possibility of this source of error in group differences does exist.

---

[9] Few of the inconsistencies represented a conflict in whether the individual had or had not participated in UB; for former nonparticipants, this was a legitimate possibility (i.e., crossovers). There were numerous cases, however, for which the particular pattern reported differed from that reported in the First Follow-up Study.

B. Establishing Causation

A logical error which is made with sufficient frequency to warrant discussion is that of inferring causality from an established relationship. While the existence of a relationship between two variables is a necessary condition for the existence of causation, it is not a sufficient one. In addition to a relationship between two variables (conditions), establishing causation requires that: (1) the causal variable (antecedent condition) logically precedes the effect (consequent condition); and (2) no other antecedent variable exists which may explain both the antecedent and consequent variables under consideration. While the first of these requirements is often fairly easy to establish, the second is extremely difficult to establish outside the true experimental paradigm. Although advances have been made recently in the area of causal modeling for nonexperimental or quasiexperimental research paradigms, the assumption of inclusion of <u>all</u> relevant variables in the model is a basic tenet in these approaches.

For educational intervention programs, it is extremely difficult, if at all possible, to identify (much less to measure) all the relevant student, process, and contextual variables which may influence one another and/or student outcomes. The UB program is no exception to this constraint. Moreover, the nature of sample selection and measurement used in the UB evaluation study was such that it was not possible in all cases to determine even the temporal order of occurrence of two events. Even when strong associations exist between previous UB participation and current outcomes, the possibility still remains that these associations are largely a function of other unconsidered (or unmeasured) antecedent variables which are related both to current outcomes and to selection of UB participants.

C. Limitations Concerning the UB Participation Pattern Variables

Most analyses considered pattern of participation in the UB program; this variable suffers some confounding in relationship to other variables. Eight unique participation patterns (including nonparticipation as a pattern) were possible in all. The eight patterns are indicated in Figure H.1. Two separate classifications of UB participation pattern were examined; one yielding two groups of former participants (typical and atypical) plus a group of nonparticipants, the other yielding four groups of former participants in addition to the nonparticipant group. The relationship of the unique patterns to the first classification is indicated by letters in Figure H.1 (NP = nonparticipant; A = atypical participant; and B = typical participant). The second classification is indicated in Figure H.1 by circled numbers. It is obvious that both of the considered classifications necessarily involved collapsing of unique participation pattern groups; as a result, some heterogeneity of unique patterns exists within the two partitionings considered for analysis.

More importantly, some participation patterns are confounded with grade cohort (e.g., grade level in fall of 1973), in that not all unique patterns are possible in a given grade. Impossible patterns are indicated by an X in Figure H.1. As an example, an individual selected as a participant in the 10th grade could legitimately exhibit only patterns including 10th grade participation: (1) 10th grade only, (2) 10th and 11th grade, (3) 10th and 12th grade, (4) 10th, 11th, and 12th grade. Similar confounds exist for other cohorts. Even though the approach taken in this study allowed for nonparticipant crossovers, which relaxes some of the constraints on possible participation

	Unique Participation Pattern							
	No Participation	Tenth Grade Only	Eleventh Grade Only	Twelfth Grade Only	Tenth and Eleventh Grades	Tenth and Twelfth Grades	Eleventh and Twelfth Grades	Tenth, Eleventh, and Twelfth Grades
Grade 10	NP 1	A 2			A 2	A 3		T 5
Grade 11	NP 1		A 2		A 2		T 4	T 5
Grade 12	NP 1			A 3		A 3	T 4	T 5

Figure H.1. Confounding of Upward Bound participation pattern and grade cohort.

pattern within a given cohort, the number of crossovers was so small that any unique pattern defined exclusively by crossovers typically had to be collapsed into another pattern classification.

Because of the constrained nature of participation pattern as a function of original cohort, analyses that did not also consider cohort may be misleading. Even though some statistical control for cohort was exercised in such analyses, the meaning of UB participation category varies as a function of cohort, suggesting an interaction, and marginal adjustment for cohort does not account for that particular confound.

One very disturbing problem with the participation pattern by grade confound was a relatively large group (N = 233) of former participants who were initially selected into the twelfth-grade cohort but who reported no twelfth-grade participation in the UB program, which is not a legitimate pattern. Only a small percentage of this group could be attributed to individuals with prior UB participation who were inappropriately selected into the nonparticipant sample. Given the sample selection process (see Appendix C), these individuals may represent those who were in the program during the summer prior to their senior year (and were still on the roster when the sample was drawn), but who did not continue their participation during their senior year. They may also represent individuals who did not report their twelfth grade participation for one reason or another. In either case, however, they represent an unusual group.

An additional problem with the UB participation pattern variable is associated with the temporal relationship between participation pattern and educational progress/persistence. While different patterns of participation in the program may be seen as theoretically preceding further educational progress, this is not necessarily the case. In particular, dropping out of high school (cessation of educational progress) leads to termination (and thus limitation) of program participation. This would likely produce over-estimates of relationships of participation pattern to educational outcomes. For this reason, examination of educational outcomes beyond high school graduation (e.g., entry into postsecondary education) was restricted to high school graduates. While this reduced the severity of the problem, it did not completely eliminate it. For example, a high school "stop out" could ultimately complete high school at a later time; the "stop out" that contributed to the atypical UB participation pattern would also introduce a time lag that would be reflected in subsequent educational progress.

Another type of confounding is related to the so-called "survivor effect," which is a form of group self-selection. As an example, consider the comparison of the educational progress of two groups of twelfth graders, one group consisting of individuals who joined UB for the first time during their senior year and the other consisting of those who participated in UB in the tenth, eleventh, and twelfth grades. The groups being compared would be basically different (even if equated on certain background factors), since one is comparing students who "survived" on their own in a school system through which the other students had been specifically assisted (i.e., to become seniors). If this survivor effect is large, then valid comparisons could only be made among those joining the program in the same grade or after controlling for the effect. Such control has not been accomplished in this study except within the tenth-grade cohort.[10] There is an obvious interaction between this confound and that discussed in the previous paragraph.

While the statistical approach and methodology employed represented an attempt to reduce, to the extent possible, problems inherent in the use of UB participation pattern as the major explanatory variable for the study, complete control was not realized, as indicated in the preceding discussion. Associated limitations in the analytic results should be anticipated.

D. Limitations of Adjustments

The approach used for these analyses relies heavily on the efficacy of the statistical adjustments, and is thus bound to the assumptions of those techniques. Although statistical adjustments were effected for the set of available measures, there is no guarantee of control over other basic differences between the participant and nonparticipant group. The use of single-degree-of-freedom contrasts, as opposed to omnibus tests, also has disadvantages in that a priori contrasts may not fully exploit the nature of the available data, and may thus obscure other results of interest, particularly interactions.

Complex statistical adjustments were performed in conducting the comparative/descriptive anslyses, and, in general, the computed standard errors compensated for the assumptions of the adjustment techniques. Nevertheless, the extent to which adjustments could be made were limited. Two basic types of adjustment were performed: (1) weight adjustments for nonresponse and item indeterminacy, and (2) statistical adjustments for the effects of variables related to both UB participation pattern and to an outcome under consideration.

Weight adjustments, in general, were "weighting class" adjustments, consisting of apportioning the sampling weights of individuals with indeterminate data to responding individuals "most like" the nonrespondent (i.e., in the same weighting class). This technique assumes that most of the nonresponse bias is captured within the categories defining "most like" (i.e., the weighting class). This assumption is more reasonable for some types of nonresponse than others; however, nonrespondents differ even from similar respondents by virtue of their nonresponse.

Weight adjustment cannot be expected to eliminate all nonresponse bias, however, except under the most trivial circumstances. The group of nonrespondents considered most problematic for this study are those who could not be located (representing sample attrition). Intuitively, less educational progress would be expected from this group than from those with whom contact was maintained. Further, the weight smoothing technique used to compensate for the nonparticipant crossover greatly increased the total weights of the former participant group (although weights for any individual originally selected as a participant relative to the total group remained stable). For this reason, estimates of total counts within the former participant groups are not considered valid. As a result of these problems, estimates of educational progress may be artifactually high.

[10] Strictly speaking, this confound is also present for comparisons between UB participants and nonparticipants. In this respect, the Markov model approach used in the base-year study was theoretically more appropriate, despite its acknowledged weaknesses.

Statistical control was also implemented for differences among former participants and nonparticipants on background and contextual factors thought to be related to educational attainment or success. Significant differences among the groups in these areas were documented in the base-year and first follow-up studies and are shown in Table H.1 (Section IV). The statistical adjustment techniques used in the second follow-up--consisting of covariance adjustment and partitioning of sample groups--introduced some problems. For one, the available analytic methodology in combination with the original sampling procedures placed stringent limits on the number of variables for which covariance adjustments could be made. Since the procedure uses within stratum estimates for the error term and since only 27 strata were defined, only 27 degrees of freedom remain to accomodate both the contrast of interest and control variables. Additionally, base rates for some variables severely limited the extent to which partitioning could be employed, while maintaining adequate cell sizes. Even greater constraints on cell sizes than would normally be imposed were introduced for appropriate calculations of standard errors, which relied on within strata estimates.

Two basic compromises were made in performing the statistical adjustments. First, covariance adjustments were restricted to marginal distributional properties; joint distributional properties (which were considered in prior studies and which allow for variable interaction) were not considered. Second, partitioning was much more restricted than would have been desired. Thus, while statistical adjustment was introduced for more variables than had previously been possible, control was not realized for the full set of potential moderating variables, and, for some factors, control was not realized to the same extent as had been achieved in previous studies. Consequently, the potential for attributing effects of some uncontrolled or only partially controlled variable to UB participation pattern is high. Although this limitation is characteristic of nonexperimental research, it should be kept in mind when considering the results.

III. ANALYTIC APPROACH

The basic analytic approach used for the comparative/descriptive analyses was specified jointly by RTI and OPBE during an early stage of the study.[11] These specifications were intended to facilitate actual conduct of analysis but were not intended to constrain investigations; thus, some deviations from the original plan were realized once data were in hand and analysis had begun. Analysis was constrained, however, by the general design of the base-year study (see Appendix C), by the set of data collected (see Appendix E), and by the specific analysis questions to be addressed by the study.

A. Basic Approach

Due to USOE's desire to introduce statistical control for pattern of UB participation and type of PSE institution attended, the balancing techniques developed in conjunction with analyses performed previously were no longer applicable owing to limitations in the original sample design. Thus, an alternate approach to adjustment was adopted with the approval of OED. The

[11] Burkheimer, G. J., Fuentes, E. J., Riccobono, J. A., and Wisenbaker, J. M. Data Analysis Plan for the Second Follow-up Study of the Upward Bound Program. Research Triangle Park, N.C.: Center for Educational Research and Evaluation, Research Triangle Institute, March 1979.

basic features of this alternate approach and its use in producing tabular results for the comparative/descriptive analyses are considered in this subsection.

Fundamentally, the approach that was employed is based upon regression analysis, although it supported the development of tabular results. Within a particular partition[12] of sample members, the general model for some outcome variable, y, for some individual, k, may be expressed as follows:

$$y_k = \mu_y + f(v_1, \ldots, v_i)_k + g(x_1, \ldots, x_j)_k + \varepsilon_k$$

where

$(v_1, \ldots, v_i)_k$ represent the values, for individual k, of coded variables based on the particular categorization of UB participation pattern under consideration,

$(x_1, \ldots, x_j)_k$ represent the values, for individual k, of control variables associated with the outcome of interest,

$f(\cdot)$ and $g(\cdot)$ are least squares linear combinations of the two sets of variables,

μ_y represents the grand mean for the outcome variable of interest,

ε_k represents the error in predicting y_k.

It should be noted that the model allows for no interaction terms within the set of control variables $(x_1, \ldots, x_j)$ or between these variables and UB participation pattern. Interactions involving the partitioning variable(s) were allowed, of course, since separate analyses were performed within each partition.

Initially, $(v_1, \ldots, v_i)$ was parameterized so as to capture the difference between a base UB participation group of interest and each of the other groups in the same partition (defining a unique value of a control variable). The particular coding used was dummy coding, which for ℓ levels of the UB participation pattern categorization (i.e, ℓ = 3 or 5 as suggested above in Figure H.1) produces ℓ-1 dichotomous (0,1) variables. Each of the ℓ-1 dummy variables results from coding one UB participation pattern group as 1 and all remaining groups as 0. The base group is the only group that is coded 0 for all ℓ-1 dummy variables. For the three-valued participation pattern categorization, "typical" participants were used as the base group; for the five-valued participation pattern categorization the base group was those who participated in grades 10, 11, and 12 (pattern 5 of Figure H.1).

Under this parameterization, the estimated regression coefficients associated with $(v_1, \ldots, v_{\ell-i})$ serve as an estimator of the difference between

---

12 Recall that additional statistical control was gained by analyzing relationships within separate partitions of the sample. Partitions were based on values of the variable(s) for which control was to be realized.

the mean of a particular participation pattern group and that of the base group; the standard error of the regression coefficient estimates the standard error of the difference between these two groups. Since the coefficients for the control variables ($x_1, \ldots, x_j$) are estimated simultaneously, the coefficients associated with ($v_1, \ldots, v_{\ell-i}$) actually reflect adjusted group differences, holding the marginal effects of other variables constant.

This information was incorporated into the tables of the comparative/descriptive analysis to produce estimated cell means adjusted for the control variables of interest (including the partitioning variable defining the rows of the table). This was accomplished by first computing the actual mean of the specified base participation pattern group in the row under consideration; that value is reported in the table. For each other cell in that particular row, two cell entries appear: the actual mean of the group, and an estimated (adjusted) mean obtained by adding the adjusted mean difference (between the group of interest and the base group) to the actual mean of the base group.

For analyses involving comparisons among the three-valued UB participation pattern categorization, the dummy coding approach, specified above, provided a parameterization that was sufficient for both producing table entries and testing specific contrasts desired by USOE. The contrasts involved differences between typical participants and (1) atypical participants, and (2) nonparticipants. For the five-valued UB participation pattern categorization, specific contrasts desired by USOE did not correspond completely with those employed to generate adjusted cell means. Moreover, when grade cohort was considered as a partitioning variable, the particular contrasts within one grade were not the same as those within another grade due to the confounding of participation pattern and grade cohort (see Table H.1).

To achieve the desired contrasts (contrasts considered in each analysis are indicated for each table in Section IV), alternative coding (difference contrasts) as required. The alternate coding produced an additional set of regression weights (appropriate to the specified contrasts) for the alternate parameterization ($v'_1, \ldots, v'_{\ell-1}$). In producing the alternate parameterization, care was taken to assure that the full set of contrasts (including those of interest) were orthogonal. The estimated standard errors associated with the solution for the alternate parameterization served as the standard errors for the contrasts of interest, and the estimated regression weights provided the point estimate of the difference contrast.

Due to the complex probability sampling employed in the selection of respondents in the base year study, standard regression analysis using ordinary least squares estimation was inappropriate. While a weighted least squares approach employing sampling weights (or adjusted sampling weights), could have provided appropriate point estimates, the associated standard errors from such an approach would not have reflected the design effects associated with the original stratification and clustering of the respondents and may have been spuriously large or small, depending on the magnitude of that design effect.

RTI had previously developed an approach to regression analysis appropriate for use with data such as these, which arise from complex sampling designs, and had produced software to implement the approach. In using this approach, which estimates error from within stratum variability, the degrees of freedom available for analysis are constrained by the number of strata and primary sampling units in the design (see below, Subsection III.B). In the present instance there are 27 degrees of freedom with which to operate. The implication for analysis was that the most complex model estimated could have no more than 26 parameters. Since each variable considered in adjustment counts as at least one parameter[13] to be estimated, relatively few control variables could be considered, even after restricting adjustment to the marginal effects for the control variables present.

With respect to the particular tables produced in the present study, this constraint was not generally a problem. Tables had at most 15 cells; the contrasts therefrom would consume no more than 14 degrees of freedom. Sufficient degrees of freedom remained for inclusion of the control variables of interest to USOE (specifically, race, sex, poverty status, academic risk status, grade cohort, and age within grade).

A particularly difficult situation existed in dealing with the case in which a table represented displays of column proportions (e.g., the distributional characteristics of a categorical outcome variable within specific participation pattern groups). The approach used in this case was to perform j separate analyses in which a dichotomous variable was predicted (e.g., separate analysis of column proportions for each of the j rows of the table). The value of the variable for the subjects in the i^{th} row was 1 if a subject was classified in level i and a 0 if not. While this yielded j separate--but not independent--analyses for each of these tables, it was the only readily-available approach whereby the control variables could be incorporated while generating approximately appropriate standard errors.

B. Taylor Series Linearization for Regression Coefficients

1. Population Definitions

Consider a finite universe U with N units U(i). Associate with the i-th unit a variate value Y(i) and a row vector of (p+1) regressors

$$X(i) = \langle X_o(i), X_1(i), \ldots, X_p(i)\rangle .$$

The linear prediction equation for Y(i) of the form

$$y(i) = X(i)\beta,$$

which minimizes the sum of squared deviations

$$Q^2 = \sum_{i \varepsilon U} [Y(i) - X(i)\beta]^2,$$

is the familiar least-squares regression equation where β is a solution to the

[13] Continuous variables generate one degree of freedom; however, categorical variables (most of the variables to be considered) generate k-1 degrees of freedom, where k is the number of unique categories defined by the variable.

so-called 'normal equations'

$$(X^TX)\beta = X^TY, \tag{H.1}$$

where

$$X^TX \equiv \sum_{i\varepsilon U} X^T(i)X(i),$$

and

$$X^TY \equiv \sum_{i\varepsilon U} X^T(i)Y(i).$$

If X^TX has rank p+1, there is a unique solution for β in H.1; namely

$$\beta = (X^TX)^{-1} X^TY.$$

If the p+1 equations represented by the matrix equation H.1 are not linearly independent, redundant equations can be replaced by independent linear restrictions on the elements of β. While the following development is limited to the full rank case, it is not difficult to extend the results directly to a particular restricted solution.

2. Estimation

Suppose that a sample S of n units is selected from the universe U. Let $\pi(i)$ denote the probability that unit U(i) will be included in such a sample. Unbiased Horvitz-Thompson estimators for X^TX and X^TY are

$$(x^Tx) = \sum_{k\varepsilon S} X^T(k)\ X(k)\ /\pi(k),$$

and

$$(x^Ty) = \sum_{k\varepsilon S} X^T(k)\ Y(k)\ /\pi(k).$$

Using these estimators, we solve for b in the estimated set of normal equations, (H.1)

$$(x^Tx)b = (x^Ty).$$

As an esimator of β, the population vector of regression coefficients, $b = (x^Tx)^{-1}\ (x^Ty)$ can be viewed as a matrix version of the combined ratio estimator $\hat{R} = (\hat{X})^{-1}(\hat{Y})$. This analogy will be strengthened by the form of the Taylor series approximation for $(b-\beta)$.

3. Taylor Series Variance Approximation

To generate the first order Taylor series approximation for b, one begins by evaluating the partial derivatives

$$\partial b/\partial(x^T y)_j, \qquad \text{for } j = 0, 1, \ldots p,$$

and

$$\partial b/\partial(x^T x)_{jj'}, \qquad \text{for } j = 0, 1, \ldots p, \quad j' = j, j+1, \ldots p,$$

where $(x^T y)_j$ represents the j-th element in the (p+1) x 1 column vector, $(x^T y)$ and $(x^T x)_{jj'}$ is the (jj')-th element of the (p+1) x (p+1) symmetric matrix $(x^T x)$.

First, notice that

$$\{\partial(x^T x)b/\partial(x^T y)_j\} = (x^T x)\ \{\partial b/\partial(x^T y)_j\}$$
$$= \partial(x^T y)/\partial(x^T y)_j.$$

Therefore

$$(x^T x)\ \{\partial b/\partial(x^T y)_j\} = \delta_j,$$

where

$$\delta_j(r) = \begin{cases} 1 \text{ if } j = r; \\ 0 \text{ otherwise.} \end{cases}$$

This allows one to write

$$\{\partial b/\partial(x^T y)_j\} = (x^T x)^{-1}\delta_j. \tag{H.2}$$

Notice that δ_j is a (p+1) x 1 column vector with a 1 in row j and zeros elsewhere. To evaluate the partial derivative of b with respect to the elements of $(X^T X)$ we note that

$$\{\partial(x^T x)b/\partial(x^T x)_{jj'}\} = \partial(x^T y)/\partial(x^T x)_{jj'} = 0.$$

Hence,

$$(x^T x)\ \{\partial b/\partial(x^T x)_{jj'}\} + \{\partial(x^T x)/\partial(x^T x)_{jj'}\}\ b = 0.$$

Recalling that $(X^T X)$ is symmetric, observe that

$$\{\partial(x^T x)/\partial(x^T x)_{jj'}\} = D_{jj'},$$

where the (r,c)-th element of $D_{jj'}$ is

$$d_{jj'}(rc) = [1 - \delta(jj')]\ \delta_j(r)\delta_{j'}(c) + \delta_{j'}(r)\delta_j(c),$$

with $\delta(jj') = 1$ if $j = j'$ and zero otherwise. This leads to

$$\{\partial b/\partial(x^Tx)_{jj'}\} = -\ (x^Tx)^{-1}D_{jj'}b, \tag{H.3}$$

where $D_{jj'}$ is a (p+1) x (p+1) symmetric matrix having 1s in positions (jj') and j'j) and zeros elsewhere.

Evaluating the partial derivatives in equations H.2 and H.3 at the point $<(x^Tx) = X^TX;\ (x^Ty) = X^TY>$, b may be approximated with the first order Taylor series linearization

$$b \doteq \beta + \sum_{j=0}^{p} [(x^Ty)_j - (X^TY_j)]\{\partial b/(x^Ty)_j\}$$

$$+ \sum_{j=0}^{p} \sum_{j'=j}^{p} [(x^Tx)_{jj'} - (X^TX)_{jj'}]\ \{\partial b/\partial(x^Tx)_{jj'}\},$$

which becomes

$$b \doteq \beta + (X^TX)^{-1} \sum_{j=0}^{p} [(x^Ty)_j - (X^TY)_j]\ \delta_j$$

$$- (X^TX)^{-1} \sum_{j=0}^{p} \sum_{j'=j}^{p} [(x^Tx)_{jj'} - (X^TX)_{jj'}]D_{jj'}\beta. \tag{H.4}$$

Recalling the definitions of δ_j and $D_{jj'}$ it is easy to see that

$$\sum_{j=0}^{p} [(x^Ty)_j - (X^TY)_j]\ \delta_j = [(x^Ty) - (X^TY)],$$

and

$$\sum_{j=0}^{p} \sum_{j'=j}^{p} [(x^Tx)_{jj'} - (X^TX)_{jj'}]\ D_{jj'} = [(x^Tx) - (X^TX)].$$

Thus, equation H.4 may be rewritten as

$$b \doteq \beta + (X^TX)^{-1} \{[x^Ty) - (X^TY)] - (x^Tx) - (X^TX)]\beta\}.$$

Finally, observing that $(X^TX)\beta = (X^TY)$,

$$b \doteq \beta + (X^TX)^{-1} [(x^Ty) - (x^Tx)\beta]. \tag{H.5}$$

To exploit the result in H.5, in order to approximate the sampling variance of b, define

$$\mathfrak{z}(k) = (X^TX)^{-1} [(x^Ty)_k - (x^Tx)_k\beta], \tag{H.6}$$

where

$$(x^Ty)_k = X^T(k)\ Y(k),$$

and

$$(x^Tx)_k = X^T(k)\ X(k).$$

The corresponding Horvitz-Thompson estimator is

$$\begin{aligned} \hat{\mathfrak{z}} &= \sum_{k \varepsilon S} \mathfrak{z}(k)/\pi(k) \\ &= (X^TX)^{-1} [(x^Ty) - (x^Tx)\beta]. \end{aligned} \tag{H.7}$$

Combining H.5 and H.7, it may be observed that

$$\hat{\mathfrak{z}} \doteq (b - \beta). \tag{H.8}$$

The result in H.8 leads one to the following approximation for the generalized mean-squared-error of b, the vector of p + 1 estimated regression coefficients:

$$\begin{aligned} \mathrm{GMSE}\{b\} &= E_S\{(b-\beta)(b-\beta)^T\} \\ &\doteq E_S\{\hat{\mathfrak{z}}\ \hat{\mathfrak{z}}^T\} \\ &= \mathrm{VAR}\{\hat{Z}\}, \end{aligned}$$

since $E_S\{\hat{z}\} = \emptyset$, the (p+1) x 1 null vector.

If $VAR_D(\hat{T})$ is the sampling variance formula for an estimated total $\hat{T}$ from a particular sampling design (D), then the generalized mean-squared-error for a vector of (p+1) regression coefficients b estimated from D is approximately

$$GMSE\{b\} \doteq VAR_D\{\hat{z}\}, \tag{H.9}$$

where $\hat{z}$ is the linearized statistic

$$\hat{z} = (X^T X)^{-1} \sum_{k \varepsilon S} [(x^T y)_k - (x^T x)_k \beta] / \pi(k).$$

In the following section, equation H.9 will be exploited to produce the 'Taylor Series' variance estimator or, more precisely, the generalized mean-squared-error estimator for b.

4. Taylor Series Variance Estimator

Recalling the Taylor series approximation for GMSE{b} developed in the previous section, it is clear that if $VAR_D(\hat{T})$ represents an appropriate variance estimator for the sample total $\hat{T}$ from design D, then $VAR_D(\hat{z})$ is the associated estimator for GMSE{b}. Since the linearized variate value $z(k)$ in H.6 is a function of the unknown population quantities $(X^T X)^{-1}$ and β, one is obliged to impose another level of approximation at this point. Instead of $z(k)$, one may use

$$z(k) = (x^T x)^{-1} [(x^T y)_k - (x^T x)_k\ b], \tag{H.10}$$

substituting the sample estimates for the unknown population parameters. It is interesting to note at this point that H.10 is a matrix analogue of the 'Taylorized' variate used to approximate the variance of the ratio $\hat{R} = (\hat{Y}/\hat{X})$. Expressing this ratio as $\hat{R} = (\hat{X})^{-1}(\hat{Y})$ and making the associations $(\hat{X})^{-1} \Leftrightarrow (x^T x)^{-1}$ and $(\hat{Y}) \Leftrightarrow (x^T y)$, the relationship between H.10 and the familiar

$$\begin{aligned} z_{\hat{R}}(k) &= [Y(k) - \hat{R}\, X\,(k)] / \hat{X} \\ &= (\hat{X})^{-1} [Y(k) - X(k)\, \hat{R}] \end{aligned}$$

is obvious.

To illustrate the method, consider a stratified simple random cluster sample with h=1(1)H strata, and n(h) clusters selected from N(h) without replacement. If y(hik) is the variate value associated with the k-th unit [k=1(1) M(hi)] in cluster-i of stratum-h and X(hik) is the corresponding 1 x (p+1) row vector of regressors, the expanded cluster totals

$$(x^T y)_{hi} = N(h) \sum_{h=1}^{M(hi)} X^T(hik)Y(hik)/n(h),$$

and

$$(x^T x)_{hi} = N(h) \sum_{k=1}^{M(hi)} X^T(hik)X(hik)/n(h),$$

may be used to form

$$\tilde{z}(hi) = [(x^T y)_{hi} - (x^T x)_{hi}\ b],$$

where

$$b = (x^T x)^{-1}\ (x^T y),$$

and

$$(x^T x) = \sum_{h=1}^{H} \sum_{i-1}^{n(h)} (x^T x)_{hi}, \text{ and}$$

$$(x^T y) = \sum_{h=1}^{H} \sum_{i=1}^{n(h)} (x^T y)_{hi}.$$

Then, GMSE{b} may be calculated as

$$\text{GMSE}\{b\} = (x^T x)^{-1} \{ \sum_{h=1}^{H} \{1 - f(h)\}\ n(h)\ s^2_{\tilde{z}}(h)\}\ (x^T x)^{-1},$$

where

$$f(h) = n(h)/N(h)$$

and

$$s^2_{\tilde{z}}(h) = \sum_{i=1}^{n(h)} [\tilde{z}(hi) - \tilde{z}(h.)][\tilde{z}(hi) - \tilde{z}(h.)]^T/[n(h)-1],$$

with $\tilde{z}(h.) = \sum_{i=1}^{n(h)} \tilde{z}(hi)/n(h)$.

5. Statistical Inference

Researchers are often interested in testing hypotheses about relationships among population regression coefficients. While there is no rigorous solution to the general linear hypothesis problem in a finite population context, a heuristic approach is suggested which relies on the central limiting tendency of estimated coefficient vectors b to have approximately the multivariate normal sampling distribution with mean β and variance-covariance matrix Var_D as specified in H.9. To the extent that this approximation for the sampling distribution of b holds, one can justify the following approach for testing

$$H_o: C^T\beta = \delta \text{ versus } H_A: C^T\beta \neq \delta;$$

form the test statistic

$$T_c^2 = (C^T b - \delta)^T [C^T \, var_D(\hat{Z}) \, C]^{-1} (C^T b - \delta), \qquad \text{(H.11)}$$

and reject H_o if T_c^2 exceeds the upper α percentage point of the Chi-square distribution with c = rank (C) degrees of freedom. this test is the multivariate analogue of the common large sample normal theory test.

When the degrees of freedom (df) associated with the estimated variance-covariance matrix $var_D(\hat{Z})$ drops below 60, H.11 may be viewed as the multivariate analogue of Student's t; namely, Hotelling's T^2 statistic. The F-transformed version of Hotelling's T^2 is expressed as

$$F = \left\{\frac{df + 1 - c}{(df)c}\right\} \quad T_c^2,$$

where df is the degrees of freedom associated with $var_D(\hat{Z})$. For the example in the previous subsection a recommended approximation is df $\doteq$ n(+)-H, where n(+) is the total number of clusters in the sample and H is the number of strata. The transformed statistic F is compared to the upper α percentage point of Fisher's F distribution with c and (df + 1 - c) degrees of freedom.

IV. TABULAR SUMMARIES OF ANALYSES

The comparative/descriptive analyses (and results) are summarized in tabular form in this section. The tables, which constitute the remainder of this appendix, have been annotated sufficiently to be interpreted without reference to textual material. Certain variables were trimmed or otherwise transformed prior to analysis; where such transformations were performed, the nature of the transformation is indicated in a footnote to the table.

Table H.1

Demographic Characteristics of the Three Major Upward Bound Participation Pattern Groups

Variable[a]	Classification	Upward Bound Participation Pattern[b] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[c]
Sex[d]	Male	.445	.403	.535	(1) NS
	Female	.555	.597	.465	(2) $p < .05$
Age Within Grade[d]	Not Over Age In Grade	.900	.871	.889	(1) NS
	Over Age In Grade	.100	.129	.111	(2) NS
Poverty Status[d]	Poverty Level	.757	.708	.354	(1) $p < .05$
	Not Poverty Level	.243	.292	.646	(2) $p < .01$
Academic Risk Status[d]	Risk	.427	.449	.480	(1) NS
	Not Risk	.573	.551	.520	(2) NS
Race	Black	.707	.561	.295	(1) $p < .01$ (2) $p < .01$
	White	.130	.249	.580	(1) $p < .01$ (2) $p < .01$
	Other	.163	.190	.125	(1) NS (2) NS
Grade Cohort[e]	10th Grade	.132	.186	.379	(1) $p < .01$ (2) $p < .01$
	11th Grade	.366	.430	.325	(1) $p < .05$ (2) NS
	12th Grade	.502	.384	.296	(1) $p < .01$ (2) $p < .01$
Number of Cases		1,357	768	1,230	
Weight Total		87,196	56,792	1,284,571	

NOTE: Cell entries are weighted column proportions within major variable categories, based on weights adjusted for indeterminacy of the variables considered and smoothed for nonparticipant crossovers.

a/ These variables determined during base year study.

b/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

c/ Unless otherwise indicated, two a priori contrasts were formed from the column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

d/ Only one set of contrasts is reported for this variable, since the contrasts for each of the two classifications of the variable are completely redundant.

e/ Represents grade level during academic year 1973-74.

Table H.2

Educational Aspiration Level
As a Function of Upward Bound Participation Pattern

Level of Aspiration	Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
	Typical	Atypical	Nonparticipant	
High School Only	(P = .016) P = .016	(P = .059) P = .033	(P = .113) P = .062	(1) NS (2) p < .05
Less Than Two Years of Vocational, Trade, or Business School	(P = .041) P = .041	(P = .054) P = .050	(P = .088) P = .086	(1) NS (2) p < .05
Two Years or More of Vocational, Trade, or Business School	(P = .056) P = .056	(P = .126) P = .127	(P = .105) P = .102	(1) p < .01 (2) p < .05
Less Than Two Years Of College	(P = .020) P = .020	(P = .027) P = .029	(P = .027) P = .034	(1) NS (2) NS
Two or More Years Of College	(P = .061) P = .061	(P = .105) P = .102	(P = .076) P = .079	(1) p < .05 (2) NS
Four- or Five-Year College Degree	(P = .258) P = .258	(P = .266) P = .275	(P = .245) P = .261	(1) NS (2) NS
Masters Degree or Equivalent	(P = .293) P = .293	(P = .191) P = .191	(P = .196) P = .188	(1) p < .01 (2) p < .01
Ph.D., M.D., or Equivalent	(P = .255) P = .255	(P = .172) P = .193	(P = .152) P = .188	(1) p < .01 (2) p < .05
Number of Cases	1,202	688	1,102	
Weight Total	87,100	56,792	1,284,667	

NOTE: Two entries are provided in each cell: (1) raw weighted within column proportions, P (in parentheses), and (2) adjusted weighted within column proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted within column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.3

Rates of Aspiring to a Four-Year Degree or More As a Function of Upward Bound Participation Pattern

Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
Typical	Atypical	Nonparticipant	
($\underset{\sim}{P}$ = .805) P = .805 N = 1,202 WN = 87,100	($\underset{\sim}{P}$ = .630) P = .660 N = 688 WN = 56,792	($\underset{\sim}{P}$ = .592) P = .637 N = 1,102 WN = 1,284,667	(1) $p < .01$ (2) $p < .01$

NOTE: Four entries are provided in each cell: (1) raw weighted proportions, $\underset{\sim}{P}$ (in parentheses), (2) adjusted weighted proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions: (1) column one versus column two, and (2) column one versus column three.

Table H.4

Educational Expectation Level
As a Function of Upward Bound Participation Pattern

Level of Expectation	Upward Bound Participation Pattern[a/] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b/]
High School Only	($\bar{P}$ = .040) P = .040	($\bar{P}$ = .135) P = .118	($\bar{P}$ = .170) P = .122	(1) p < .01 (2) p < .01
Less Than Two Years of Vocational, Trade, or Business School	($\bar{P}$ = .039) P = .039	($\bar{P}$ = .079) P = .071	($\bar{P}$ = .091) P = .097	(1) NS (2) p < .01
Two Years or More of Vocational, Trade, or Business School	($\bar{P}$ = .071) P = .071	($\bar{P}$ = .099) P = .093	($\bar{P}$ = .091) P = .071	(1) NS (2) NS
Less Than Two Years Of College	($\bar{P}$ = .035) P = .035	($\bar{P}$ = .047) P = .045	($\bar{P}$ = .071) P = .079	(1) NS (2) p < .01
Two or More Years Of College	($\bar{P}$ = .096) P = .096	($\bar{P}$ = .098) P = .091	($\bar{P}$ = .109) P = .102	(1) NS (2) NS
Four- or Five-Year College Degree	($\bar{P}$ = .342) P = .342	($\bar{P}$ = .312) P = .326	($\bar{P}$ = .257) P = .279	(1) NS (2) NS
Masters Degree or Equivalent	($\bar{P}$ = .268) P = .268	($\bar{P}$ = .161) P = .166	($\bar{P}$ = .155) P = .156	(1) p < .01 (2) p < .01
Ph.D., M.D., or Equivalent	($\bar{P}$ = .108) P = .108	($\bar{P}$ = .070) P = .090	($\bar{P}$ = .057) P = .093	(1) NS (2) NS
Number of Cases	1,096	624	962	
Weight Total	87,005	56,792	1,284,763	

NOTE: Two entries are provided in each cell: (1) raw weighted within column proportions, $\bar{P}$ (in parentheses), and (2) adjusted weighted within column proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two a priori contrasts were formed from the adjusted within column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.5

Rates of Expecting a Four-Year Degree or More
As a Function of Upward Bound Participation Pattern

Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
Typical	Atypical	Nonparticipant	
($\underset{\sim}{P}$ = .718) P = .718 N = 1,096 WN = 87,005	($\underset{\sim}{P}$ = .543) P = .583 N = 624 WN = 56,792	($\underset{\sim}{P}$ = .469) P = .529 N = 962 WN = 1,284,763	(1) $p < .01$ (2) $p < .01$

NOTE: Four entries are provided in each cell: (1) raw weighted proportions, $\underset{\sim}{P}$ (in parentheses), (2) adjusted weighted proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions: (1) column one versus column two, and (2) column one versus column three.

Table H.6

High School Completion Rates as a Function of Upward Bound Participation Pattern and Cohort

Cohort (Grade Level In Fall 1973)	Upward Bound Participation Pattern[a]				
	10th, 11th, and 12th Grade	11th and 12th Grade	12th Grade Only Or 10th and 12th Grade	10th Grade Only, 11th Grade Only, Or 10th and 11th Grade	No Participation
10th Grade[b]	(P = .987) $\bar{P}$ = .987 N = 161 WN = 11,500			(P = .905) $\bar{P}$ = .935 N = 129 WN = 10,583	(P = .927) $\bar{P}$ = .934 N = 375 WN = 486,689
11th Grade[c]	(P = .989) $\bar{P}$ = .989 N = 287 WN = 18,299	(P = .983) $\bar{P}$ = .985 N = 214 WN = 13,650		(P = .953) $\bar{P}$ = .955 N = 320 WN = 24,381	(P = .950) $\bar{P}$ = .957 N = 405 WN = 417,151
12th Grade[d]	(P = .997) $\bar{P}$ = .997 N = 355 WN = 22,688	(P = 1.000) $\bar{P}$ = 1.000 N = 338 WN = 21,060	(P = .991) $\bar{P}$ = .987 N = 81 WN = 5,361	(P = .978) $\bar{P}$ = .972 N = 233 WN = 16,467	(P = .980) $\bar{P}$ = .960 N = 442 WN = 380,731

NOTE: Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses); (2) adjusted weighted proportions, $\bar{P}$ (adjusted within row for sex, race, poverty status, academic risk status, and over age in grade characteristics of the leftmost UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Three cells in the crossing of participation pattern and cohort have been eliminated. Two such cells represent impossible situations for the original sample of UB participants (but were possible for the crossover group). The remaining cell (10th and 12th grade participation for 10th grade cohort members) represented a very atypical pattern. Individuals falling in the two eliminated cells in column three were relocated into the cells immediately to the right; those falling in the eliminated cell in column two were relocated into the cell immediately to the left.

b/ Two *a priori* contrasts were formed from the adjusted proportions of this row: (1) column one versus column four, and (2) column one versus column five. Neither contrast was statistically significant at the .05 level.

c/ Three *a priori* contrasts were formed from the adjusted proportions of this row: (1) column one versus column two, (2) column one versus column five, and (3) columns one and two versus column four. The final contrast was statistically significant at the .05 level.

d/ Four *a priori* contrasts were formed from the adjusted proportions of this row: (1) column one versus column two, (2) column one versus column five, (3) column three versus column four, and (4) columns one and two versus columns three and four. The second and fourth contrasts were statistically significant at the .05 level.

Table H.7

Postsecondary Entry Rates as a Function of Upward Bound Participation Pattern and Cohort

Cohort (Grade Level In Fall 1973)	Upward Bound Participation Pattern[a]				
	10th, 11th, and 12th Grade	11th and 12th Grade	12th Grade Only Or 10th and 12th Grade	10th Grade Only, 11th Grade Only, Or 10th and 11th Grade	No Participati
10th Grade[b]	(P = .859) $\tilde{P}$ = .859 N = 150 WN = 11,347			(P = .676) $\tilde{P}$ = .769 N = 112 WN = 9,575	(P = .649) $\tilde{P}$ = .739 N = 341 WN = 451,110
11th Grade[c]	(P = .935) $\tilde{P}$ = .935 N = 279 WN = 18,061	(P = .901) $\tilde{P}$ = .904 N = 206 WN = 13,444		(P = .718) $\tilde{P}$ = .711 N = 296 WN = 23,238	(P = .636) $\tilde{P}$ = .638 N = 376 WN = 396,193
12th Grade[d]	(P = .896) $\tilde{P}$ = .896 N = 348 WN = 22,509	(P = .918) $\tilde{P}$ = .934 N = 334 WN = 21,163	(P = .872) $\tilde{P}$ = .920 N = 76 WN = 5,195	(P = .747) $\tilde{P}$ = .791 N = 219 WN = 16,218	(P = .631) $\tilde{P}$ = .724 N = 426 WN = 373,027

NOTE: These analyses were restricted to high school graduates. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses); (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for sex, race, poverty status, academic risk status, and over age in grade characteristics of the leftmost UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Three cells in the crossing of participation pattern and cohort have been eliminated. Two such cells represent impossible situations for the original sample of UB participants (but were possible for the crossover group). The remaining cell (10th and 12th grade participation for 10th grade cohort members) represented a very atypical pattern. Individuals falling in the two eliminated cells in column three were relocated into the cells immediately to the right; those falling in the eliminated cell in column two were relocated into the cell immediately to the left.

b/ Two *a priori* contrasts were formed from the adjusted proportions of this row: (1) column one versus column four, and (2) column one versus column five. Neither contrast was statistically significant at the .05 level.

c/ Three *a priori* contrasts were formed from the adjusted proportions of this row: (1) column one versus column two, (2) column one versus column five, and (3) columns one and two versus column four. The latter two contrasts were statistically significant at the .05 level.

d/ Four *a priori* contrasts were formed from the adjusted proportions of this row: (1) column one versus column two, (2) column one versus column five, (3) column three versus column four, and (4) columns one and two versus columns three and four. Only the second contrast was statistically significant at the .05 level.

Table H.8

Rates of Lag Entry into Postsecondary Institutions as a Function of Upward Bound Participation Pattern and Cohort

Cohort (Grade Level In Fall 1973)	Upward Bound Participation Pattern[a]				
	10th, 11th, and 12th Grade	11th and 12th Grade	12th Grade Only Or 10th and 12th Grade	10th Grade Only, 11th Grade Only, Or 10th and 11th Grade	No Participation
10th Grade[b]	(P = .130) $\bar{P}$ = .130 N = 129 WN = 9,926			(P = .348) $\bar{P}$ = .353 N = 72 WN = 6,571	(P = .258) $\bar{P}$ = .338 N = 203 WN = 297,814
11th Grade[c]	(P = .153) $\bar{P}$ = .153 N = 252 WN = 17,102	(P = .104) $\bar{P}$ = .105 N = 181 WN = 12,044		(P = .247) $\bar{P}$ = .249 N = 199 WN = 16,799	(P = .170) $\bar{P}$ = .180 N = 239 WN = 252,996
12th Grade[d]	(P = .097) $\bar{P}$ = .097 N = 304 WN = 20,206	(P = .136) $\bar{P}$ = .128 N = 303 WN = 19,523	(P = .122) $\bar{P}$ = .103 N = 68 WN = 4,842	(P = .254) $\bar{P}$ = .244 N = 166 WN = 12,246	(P = .261) $\bar{P}$ = .316 N = 280 WN = 237,453

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses); (2) adjusted weighted proportions, $\bar{P}$ (adjusted within row for sex, race, poverty status, academic risk status, and over age in grade characteristics of the leftmost UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Three cells in the crossing of participation pattern and cohort have been eliminated. Two such cells represent impossible situations for the original sample of UB participants (but were possible for the crossover group). The remaining cell (10th and 12th grade participation for 10th grade cohort members) represented a very atypical pattern. Individuals falling in the two eliminated cells in column three were relocated into the cells immediately to the right; those falling in the eliminated cell in column two were relocated into the cell immediately to the left.

b/ Two *a priori* contrasts were formed from the adjusted proportions of this row: (1) column one versus column four, and (2) column one versus column five. Both contrasts were statistically significant at the .05 level.

c/ Three *a priori* contrasts were formed from the adjusted proportions of this row: (1) column one versus column two, (2) column one versus column five, and (3) columns one and two versus column four. Only the last of these contrasts was statistically significant at the .05 level.

d/ Four *a priori* contrasts were formed from the adjusted proportions of this row: (1) column one versus column two, (2) column one versus column five, (3) column three versus column four, and (4) columns one and two versus columns three and four. All but the first of these contrasts were statistically significant at the .05 level.

Table H.9

Educational Progress Relative to Goals
As a Function of Upward Bound Participation Pattern and Cohort

Cohort (Grade Level in Fall 1973)[b]	Upward Bound Participation Pattern[a]			Significance Level Of Contrasts
	Typical	Atypical	Nonparticipant	
10th Grade	(M = 40.1%) M̃ = 40.1% N = 129 WN = 11,157	(M = 23.3%) M̃ = 27.6% N = 103 WN = 9,930	(M = 30.0%) M̃ = 32.6% N = 292 WN = 430,855	(1) $p < .05$ (2) NS
11th Grade	(M = 51.8%) M̃ = 51.8% N = 428 WN = 31,320	(M = 37.7%) M̃ = 36.2% N = 263 WN = 22,927	(M = 39.2%) M̃ = 35.4% N = 324 WN = 369,861	(1) $p < .01$ (2) $p < .01$
12th Grade	(M = 52.0%) M̃ = 52.0% N = 601 WN = 43,421	(M = 44.2%) M̃ = 44.8% N = 257 WN = 20,815	(M = 37.8%) M̃ = 38.2% N = 349 WN = 349,105	(1) $p < .05$ (2) $p < .01$

NOTE: These analyses are restricted to those aspiring to some form of postsecondary education. The variable reported is actual progress expressed as a percentage of progress necessary to achieve reported educational goal (see Appendix E). Four entries are provided in each cell: (1) raw weighted means, M (in parentheses), (2) adjusted weighted means, M̃ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

[a] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[b] Two *a priori* contrasts were formed from the adjusted means within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.10

Extent to Which Poor Grades Were Reported as an Educational Problem Area
As a Function of Upward Bound Participation Pattern

Extent of Problem	Upward Bound Participation Pattern[a] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b]
Did Not Experience	(P = .352) P = .352	(P = .461) P = .446	(P = .478) P = .415	(1) $p < .01$ (2) NS
Small Problem	(P = .427) P = .427	(P = .348) P = .373	(P = .396) P = .454	(1) NS (2) NS
Moderate Problem	(P = .159) P = .159	(P = .135) P = .129	(P = .083) P = .083	(1) NS (2) $p < .01$
Large Problem	(P = .062) P = .062	(P = .056) P = .052	(P = .044) P = .048	(1) NS (2) NS
Mean Value[c]	(M = 1.93) M = 1.93	(M = 1.78) M = 1.79	(M = 1.69) M = 1.76	(1) $p < .01$ (2) $p < .01$
Number of Cases	1,221	710	1,144	
Weight Total	87,196	56,792	1,284,571	

NOTE: Two entries are provided in each cell: (1) raw weighted column proportions or means, P or M (in parentheses), and (2) adjusted weighted column proportions or means, P or M (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

[a] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[b] Two a priori contrasts were formed from the adjusted column proportions or means within each row: (1) column one versus column two, and (2) column one versus column three.

[c] Averages were taken for a four-valued scaling (from 1 to 4 in the order given) of extent to which problem was experienced.

Table H.11

Extent to Which Poor High School Preparation Was Reported as an Educational Problem Area As a Function of Upward Bound Participation Pattern

Extent of Problem	Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
	Typical	Atypical	Nonparticipant	
Did Not Experience	(P = .433) P = .433	(P = .459) P = .457	(P = .526) P = .517	(1) NS (2) $p < .05$
Small Problem	(P = .271) P = .271	(P = .287) P = .286	(P = .285) P = .282	(1) NS (2) NS
Moderate Problem	(P = .144) P = .144	(P = .130) P = .137	(P = .117) P = .124	(1) NS (2) NS
Large Problem	(P = .152) P = .152	(P = .124) P = .120	(P = .071) P = .077	(1) NS (2) $p < .01$
Mean Value[c]	(M = 2.02) M = 2.02	(M = 1.92) M = 1.92	(M = 1.73) M = 1.76	(1) NS (2) $p < .01$
Number of Cases	1,221	710	1,144	
Weight Total	87,196	56,792	1,284,571	

NOTE: Two entries are provided in each cell: (1) raw weighted column proportions or means, P or M (in parentheses), and (2) adjusted weighted column proportions or means, P or M (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions or means within each row: (1) column one versus column two, and (2) column one versus column three.

c/ Averages were taken for a four-valued scaling (from 1 to 4 in the order given) of extent to which problem was experienced.

Table H.12

Extent to Which Lack of Personal or Family Finances Was Reported as an Educational Problem Area As a Function of Upward Bound Participation Pattern

Extent of Problem	Upward Bound Participation Pattern[a] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b]
Did Not Experience	(P = .245) $\tilde{P}$ = .245	(P = .301) $\tilde{P}$ = .270	(P = .481) $\tilde{P}$ = .358	(1) NS (2) $p < .05$
Small Problem	(P = .205) $\tilde{P}$ = .205	(P = .232) $\tilde{P}$ = .243	(P = .214) $\tilde{P}$ = .231	(1) NS (2) NS
Moderate Problem	(P = .232) $\tilde{P}$ = .232	(P = .166) $\tilde{P}$ = .172	(P = .166) $\tilde{P}$ = .192	(1) $p < .01$ (2) NS
Large Problem	(P = .319) $\tilde{P}$ = .319	(P = .301) $\tilde{P}$ = .315	(P = .138) $\tilde{P}$ = .219	(1) NS (2) $p < .01$
Mean Value[c]	(M = 2.62) $\tilde{M}$ = 2.62	(M = 2.47) $\tilde{M}$ = 2.53	(M = 1.96) $\tilde{M}$ = 2.27	(1) NS (2) $p < .01$
Number of Cases	1,221	710	1,144	
Weight Total	87,196	56,792	1,284,571	

NOTE: Two entries are provided in each cell: (1) raw weighted column proportions or means, P or M (in parentheses), and (2) adjusted weighted column proportions or means, $\tilde{P}$ or $\tilde{M}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

[a] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[b] Two *a priori* contrasts were formed from the adjusted column proportions or means within each row: (1) column one versus column two, and (2) column one versus column three.

[c] Averages were taken for a four-valued scaling (from 1 to 4 in the order given) of extent to which problem was experienced.

Table H.13

Extent to Which Inadequate Financial Aid Was Reported as an Educational Problem Area As a Function of Upward Bound Participation Pattern

Extent of Problem	Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
	Typical	Atypical	Nonparticipant	
Did Not Experience	(P = .411) P = .411	(P = .444) P = .420	(P = .579) P = .475	(1) NS (2) NS
Small Problem	(P = .247) P = .247	(P = .234) P = .245	(P = .205) P = .229	(1) NS (2) NS
Moderate Problem	(P = .192) P = .192	(P = .149) P = .156	(P = .128) P = .165	(1) $p < .01$ (2) NS
Large Problem	(P = .151) P = .151	(P = .173) P = .180	(P = .088) P = .131	(1) NS (2) NS
Mean Value[c]	(M = 2.08) M = 2.08	(M = 2.05) M = 2.10	(M = 1.73) M = 1.95	(1) NS (2) NS
Number of Cases	1,221	710	1,143	
Weight Total	87,196	56,792	1,284,571	

NOTE: Two entries are provided in each cell: (1) raw weighted column proportions or means, P or M (in parentheses), and (2) adjusted weighted column proportions or means, P or M (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions or means within each row: (1) column one versus column two, and (2) column one versus column three.

c/ Averages were taken for a four-valued scaling (from 1 to 4 in the order given) of extent to which problem was experienced.

Table H.14

Extent to Which Unavailability of Desired Program Was Reported as an Educational Problem Area
As a Function of Upward Bound Participation Pattern

Extent of Problem	Upward Bound Participation Pattern[a] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b]
Did Not Experience	(P = .700) P = .700	(P = .702) P = .686	(P = .717) P = .673	(1) NS (2) NS
Small Problem	(P = .151) P = .151	(P = .176) P = .178	(P = .154) P = .155	(1) NS (2) NS
Moderate Problem	(P = .071) P = .071	(P = .058) P = .061	(P = .066) P = .069	(1) NS (2) NS
Large Problem	(P = .077) P = .077	(P = .064) P = .075	(P = .063) P = .103	(1) NS (2) NS
Mean Value[c]	(M = 1.52) M = 1.52	(M = 1.48) M = 1.52	(M = 1.48) M = 1.60	(1) NS (2) NS
Number of Cases	1,221	710	1,142	
Weight Total	87,196	56,792	1,284,571	

NOTE: Two entries are provided in each cell: (1) raw weighted column proportions or means, P or M (in parentheses), and (2) adjusted weighted column proportions or means, P or M (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions or means within each row: (1) column one versus column two, and (2) column one versus column three.

c/ Averages were taken for a four-valued scaling (from 1 to 4 in the order given) of extent to which problem was experienced.

Table H.15

Extent to Which Family Responsibilities Were Reported as an Educational Problem Area
As a Function of Upward Bound Participation Pattern

Extent of Problem	Upward Bound Participation Pattern[a/] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b/]
Did Not Experience	(P = .525) $\tilde{P}$ = .525	(P = .563) $\tilde{P}$ = .542	(P = .633) $\tilde{P}$ = .523	(1) NS (2) NS
Small Problem	(P = .232) $\tilde{P}$ = .232	(P = .206) $\tilde{P}$ = .209	(P = .203) $\tilde{P}$ = .220	(1) NS (2) NS
Moderate Problem	(P = .124) $\tilde{P}$ = .124	(P = .109) $\tilde{P}$ = .119	(P = .079) $\tilde{P}$ = .133	(1) NS (2) NS
Large Problem	(P = .119) $\tilde{P}$ = .119	(P = .122) $\tilde{P}$ = .129	(P = .084) $\tilde{P}$ = .125	(1) NS (2) NS
Mean Value[c/]	(M = 1.84) $\tilde{M}$ = 1.84	(M = 1.79) $\tilde{M}$ = 1.84	(M = 1.61) $\tilde{M}$ = 1.86	(1) NS (2) NS
Number of Cases	1,221	710	1,144	
Weight Total	87,196	56,792	1,284,571	

NOTE: Two entries are provided in each cell: (1) raw weighted column proportions or means, P or M (in parentheses), and (2) adjusted weighted column proportions or means, $\tilde{P}$ or $\tilde{M}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions or means within each row: (1) column one versus column two, and (2) column one versus column three.

c/ Averages were taken for a four-valued scaling (from 1 to 4 in the order given) of extent to which problem was experienced.

Table H.16

Extent to Which Lack of Information on Desired Program Was Reported as an Educational Problem Area As a Function of Upward Bound Participation Pattern

Extent of Problem	Upward Bound Participation Pattern[a] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b]
Did Not Experience	(P = .592) P = .592	(P = .638) P = .601	(P = .654) P = .551	(1) NS (2) NS
Small Problem	(P = .215) P = .215	(P = .198) P = .208	(P = .185) P = .213	(1) NS (2) NS
Moderate Problem	(P = .107) P = .107	(P = .087) P = .102	(P = .096) P = .135	(1) NS (2) NS
Large Problem	(P = .086) P = .086	(P = .077) P = .090	(P = .065) P = .101	(1) NS (2) NS
Mean Value[c]	(M = 1.69) M = 1.69	(M = 1.60) M = 1.68	(M = 1.57) M = 1.79	(1) NS (2) NS
Number of Cases	1,221	710	1,143	
Weight Total	87,196	56,792	1,284,571	

NOTE: Two entries are provided in each cell: (1) raw weighted column proportions or means, P or M (in parentheses), and (2) adjusted weighted column proportions or means, P or M (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

[a] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[b] Two *a priori* contrasts were formed from the adjusted column proportions or means within each row: (1) column one versus column two, and (2) column one versus column three.

[c] Averages were taken for a four-valued scaling (from 1 to 4 in the order given) of extent to which problem was experienced.

Table H.17

Extent To Which Lack of Information on Educational Requirements for Desired Career
Was Reported as an Educational Problem Area
As a Function of Upward Bound Participation Pattern

Extent of Problem	Upward Bound Participation Pattern[a/] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b/]
Did Not Experience	(P = .573) P = .573	(P = .572) P = .542	(P = .617) P = .539	(1) NS (2) NS
Small Problem	(P = .218) P = .218	(P = .239) P = .246	(P = .199) P = .226	(1) NS (2) NS
Moderate Problem	(P = .112) P = .112	(P = .084) P = .099	(P = .083) P = .115	(1) NS (2) NS
Large Problem	(P = .097) P = .097	(P = .105) P = .113	(P = .101) P = .119	(1) NS (2) NS
Mean Value[c/]	(M = 1.32) M = 1.32	(M = 1.72) M = 1.78	(M = 1.67) M = 1.81	(1) NS (2) NS
Number of Cases	1,221	710	1,143	
Weight Total	87,196	56,792	1,284,571	

NOTE: Two entries are provided in each cell: (1) raw weighted column proportions or means, P or M (in parentheses), and (2) adjusted weighted column proportions or means, P or M (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions or means within each row: (1) column one versus column two, and (2) column one versus column three.

c/ Averages were taken for a four-valued scaling (from 1 to 4 in the order given) of extent to which problem was experienced.

Table H.18

Current Employment Status
As a Function of Upward Bound Participation Pattern
And Current Enrollment Status

Current Enrollment Status	Current Employment Status	Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
		Typical	Atypical	Nonparticipant	
Not Enrolled	Not Working	(P = .271) P = .271	(P = .278) P = .256	(P = .189) P = .180	(1) NS (2) $p < .05$
	Working Part-Time	(P = .119) P = .119	(P = .139) P = .145	(P = .105) P = .128	(1) NS (2) NS
	Working Full-Time	(P = .610) P = .610	(P = .583) P = .599	(P = .706) P = .692	(1) NS (2) NS
	Number of Cases	570	249	343	
	Weight Total	43,680	29,002	590,219	
Enrolled	Not Working	(P = .454) P = .454	(P = .475) P = .495	(P = .432) P = .520	(1) NS (2) NS
	Working Part-Time	(P = .292) P = .292	(P = .203) P = .163	(P = .304) P = .218	(1) $p < .01$ (2) NS
	Working Full-Time	(P = .254) P = .254	(P = .323) P = .342	(P = .264) P = .262	(1) $p < .05$ (2) NS
	Number of Cases	541	210	343	
	Weight Total	43,362	27,790	694,507	

NOTE: Two entries are provided in each cell: (1) raw weighted column proportions within major enrollment-status breakdown, P (in parentheses), and (2) adjusted weighted column proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.19

Average Level of Satisfaction With Life
As a Function of Upward Bound Participation Pattern and Current Educational Status

Current Educational Status	Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
	Typical	Atypical	Nonparticipant	
Did Not Enter Postsecondary Education	(M = 3.13) $\tilde{M}$ = 3.13 N = 98 WN = 7,459	(M = 2.70) $\tilde{M}$ = 2.91 N = 165 WN = 14,705	(M = 3.51) $\tilde{M}$ = 3.22 N = 353 WN = 450,311	(1) NS (2) NS
Entered Postsecondary Education But Did Not Graduate and Not Currently Enrolled	(M = 2.70) $\tilde{M}$ = 2.70 N = 363 WN = 27,983	(M = 3.11) $\tilde{M}$ = 3.06 N = 152 WN = 12,532	(M = 3.44) $\tilde{M}$ = 2.84 N = 202 WN = 227,958	(1) $p < .05$ (2) NS
Entered Postsecondary Education and Graduated Or Currently Enrolled	(M = 3.60) $\tilde{M}$ = 3.60 N = 671 WN = 50,838	(M = 3.50) $\tilde{M}$ = 3.44 N = 283 WN = 26,990	(M = 3.74) $\tilde{M}$ = 3.33 N = 452 WN = 542,304	(1) NS (2) $p < .05$

NOTE: Satisfaction was measured on a five point scale from 1 = very dissatisfied to 5 = very satisfied. Four entries are provided in each cell: (1) raw weighted means, M (in parentheses), (2) adjusted weighted means, $\tilde{M}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.20

Demographic Characteristics of PSE Entrants
From the Three Major Upward Bound Participation Pattern Groups

Variable[a]	Classification	Upward Bound Participation Pattern[b]			Significance Level Of Contrasts[c]
		Typical	Atypical	Nonparticipant	
Sex[d]	Male	.430	.384	.508	(1) NS
	Female	.570	.616	.492	(2) NS
Age Within Grade[d]	Not Over Age In Grade	.906	.899	.917	(1) NS (2) NS
	Over Age In Grade	.094	.101	.083	
Poverty Status[d]	Poverty Level	.758	.714	.309	(1) NS
	Not Poverty Level	.242	.286	.691	(2) $p < .01$
Academic Risk Status[d]	Risk	.409	.390	.375	(1) NS
	Not Risk	.591	.610	.625	(2) NS
Race	Black	.724	.616	.318	(1) $p < .01$ (2) $p < .01$
	White	.118	.191	.579	(1) $p < .01$ (2) $p < .01$
	Other	.158	.193	.103	(1) NS (2) NS
Grade Cohort[e]	10th Grade	.120	.164	.370	(1) $p < .05$ (2) $p < .01$
	11th Grade	.372	.419	.329	(1) NS (2) NS
	12th Grade	.508	.417	.301	(1) $p < .05$ (2) $p < .01$
Number of Cases		1,196	515	746	
Weight Total		76,960	38,567	760,281	

NOTE: These analyses were restricted to PSE entrants. Cell entries are weighted column proportions within major variable categories, based on weights adjusted for indeterminacy of the variables considered and smoothed for nonparticipant crossovers.

a/ These variables determined during base year study.

b/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

c/ Unless otherwise indicated, two a priori contrasts were formed from the column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

d/ Only one set of contrasts is reported for this variable, since the contrasts for each of the two classifications of the variable are completely redundant.

e/ Represents grade level during academic year 1973-74.

Table H.21

Student Body Family Income Level of Institution Last Attended
As a Function of Upward Bound Participation Pattern

Family Income Level Of Student Body	Upward Bound Participation Pattern[a/] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b/]
Low SES	(P = .578) P = .578	(P = .578) P = .596	(P = .459) P = .580	(1) NS (2) NS
Intermediate SES	(P = .352) P = .352	(P = .364) P = .354	(P = .454) P = .367	(1) NS (2) NS
High SES	(P = .070) P = .070	(P = .058) P = .050	(P = .087) P = .053	(1) NS (2) NS
Number of Cases	1,171	506	736	
Weight Total	76,960	38,567	762,304	

NOTE: These analyses were restricted to PSE entrants. Two entries are provided in each cell: (1) raw weighted column proportions, P (in parentheses), and (2) adjusted weighted column proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.22

Relative Selectivity Level of Institution Last Attended
As a Function of Upward Bound Participation Pattern

Relative Selectivity Level	Upward Bound Participation Pattern[a/]			Significance Level Of Contrasts[b/]
	Typical	Atypical	Nonparticipant	
Low Selectivity	(P = .582) $\tilde{P}$ = .582	(P = .638) $\tilde{P}$ = .660	(P = .579) $\tilde{P}$ = .699	(1) NS (2) $p < .05$
Intermediate Selectivity	(P = .239) $\tilde{P}$ = .239	(P = .240) $\tilde{P}$ = .223	(P = .219) $\tilde{P}$ = .149	(1) NS (2) $p < .05$
High Selectivity	(P = .179) $\tilde{P}$ = .179	(P = .123) $\tilde{P}$ = .117	(P = .202) $\tilde{P}$ = .152	(1) $p < .05$ (2) NS
Number of Cases	1,171	506	736	
Weight Total	76,960	38,567	762,304	

NOTE: These analyses were restricted to PSE entrants. Two entries are provided in each cell: (1) raw weighted column proportions, P (in parentheses), and (2) adjusted weighted column proportions, $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.23

Distribution of Cost of Institution Last Attended
As a Function of Upward Bound Participation Pattern

Composite Per-Student Cost of Attending Institution	Upward Bound Participation Pattern[a/] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b/]
Less Than $500	(P = .294) P = .294	(P = .338) P = .318	(P = .404) P = .379	(1) NS (2) NS
$500 - $999	(P = .407) P = .407	(P = .384) P = .402	(P = .420) P = .455	(1) NS (2) NS
$1,000 - $1,999	(P = .140) P = .140	(P = .112) P = .119	(P = .086) P = .098	(1) NS (2) NS
$2,000 - $3,999	(P = .132) P = .132	(P = .145) P = .145	(P = .071) P = .070	(1) NS (2) NS
$4,000 or More	(P = .028) P = .028	(P = .022) P = .016	(P = .020) P = .003	(1) NS (2) $p < .05$
Number of Cases	1,159	487	711	
Weight Total	76,960	38,567	762,304	

NOTE: These analyses were restricted to PSE entrants. Two entries are provided in each cell: (1) raw weighted column proportions, P (in parentheses), and (2) adjusted weighted column proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.24

Size of Institution Last Attended
As a Function of Upward Bound Participation Pattern

Undergradate Student Enrollment	Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
	Typical	Atypical	Nonparticipant	
Less Than 500	(P = .086) P = .086	(P = .182) P = .187	(P = .156) P = .219	(1) $p < .01$ (2) $p < .01$
500 - 2,499	(P = .358) P = .358	(P = .317) P = .309	(P = .220) P = .206	(1) NS (2) $p < .01$
2,500 - 7,499	(P = .273) P = .273	(P = .241) P = .245	(P = .260) P = .224	(1) NS (2) NS
7,500 - 14,999	(P = .167) P = .167	(P = .153) P = .152	(P = .226) P = .191	(1) NS (2) NS
15,000 or More	(P = .115) P = .115	(P = .107) P = .107	(P = .138) P = .160	(1) NS (2) NS
Number of Cases	1,171	506	736	
Weight Total	76,960	38,567	762,304	

NOTE: These analyses were restricted to PSE entrants. Two entries are provided in each cell: (1) raw weighted column proportions, P (in parentheses), and (2) adjusted weighted column proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

[a] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[b] Two a priori contrasts were formed from the adjusted column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.25

Rates of Attending Institutions
With Predominantly or Traditionally Non-White Student Bodies
As a Function of Upward Bound Participation

Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
Typical	Atypical	Nonparticipant	
($\underset{\sim}{P}$ = .297)	($\underset{\sim}{P}$ = .170)	($\underset{\sim}{P}$ = .117)	(1) $p < .01$
$\tilde{P}$ = .297	$\tilde{P}$ = .172	$\tilde{P}$ = .179	(2) $p < .01$
N = 1,171	N = 506	N = 736	
WN = 76,960	WN = 38,567	WN = 762,304	

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, $\underset{\sim}{P}$ (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.26

Offering and Control Level of Institution Last Attended
As a Function of Upward Bound Participation Pattern

Offering/Control Classification	Upward Bound Participation Pattern[a/]			Significance Level Of Contrasts[b/]
	Typical	Atypical	Nonparticipant	
Vocational, Technical, Business, or Proprietary School	(P = .084) P = .084	(P = .181) P = .188	(P = .172) P = .221	(1) p < .01 (2) p < .01
Public Two-Year College	(P = .181) P = .181	(P = .248) P = .244	(P = .255) P = .267	(1) p < .05 (2) NS
Private Two-Year College	(P = .009) P = .009	(P = .024) P = .024	(P = .017) P = .016	(1) NS (2) NS
Public Four-Year College Or University	(P = .460) P = .460	(P = .360) P = .355	(P = .413) P = .357	(1) p < .01 (2) p < .05
Private Four-Year College Or University	(P = .266) P = .266	(P = .187) P = .190	(P = .143) P = .139	(1) p < .05 (2) p < .01
Number of Cases	1,171	506	736	
Weight Total	76,960	38,567	762,304	

NOTE: These analyses were restricted to PSE entrants. Two entries are provided in each cell: (1) raw weighted column proportions, P (in parentheses), and (2) adjusted weighted column proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two a priori contrasts were formed from the adjusted column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.27

Proportion of Two-Year Postsecondary Institution Entrants
Subsequently Transferring to Four-Year Institutions
As a Function of Upward Bound Participation Pattern

Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
Typical	Atypical	Nonparticipant	
(P = .217)	(P = .199)	(P = .173)	(1) NS
$\tilde{P}$ = .217	$\tilde{P}$ = .190	$\tilde{P}$ = .160	(2) NS
N = 434	N = 278	N = 454	
WN = 28,644	WN = 21,395	WN = 439,697	

NOTE: These analyses were restricted to those who entered two-year postsecondary institutions. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell size, WN.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions: (1) column one versus column two, and (2) column one versus column three.

Table H.28

Rates of Attending Upward Bound Host Institutions
As a Function of Upward Bound Participation Pattern

Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
Typical	Atypical	Nonparticipant	
(P = .533)	(P = .347)	(P = .317)	(1) $p < .01$
$\tilde{P}$ = .533	$\tilde{P}$ = .354	$\tilde{P}$ = .359	(2) $p < .01$
N = 1,171	N = 506	N = 736	
WN = 76,960	WN = 38,567	WN = 762,304	

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions: (1) column one versus column two, and (2) column one versus column three.

Table H.29

Rates of Attending Special Services Host Institutions
As a Function of Upward Bound Participation Pattern

Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
Typical	Atypical	Nonparticipant	
(P = .546) $\tilde{P}$ = .546 N = 1,171 WN = 76,960	(P = .411) $\tilde{P}$ = .409 N = 506 WN = 38,567	(P = .474) $\tilde{P}$ = .472 N = 736 WN = 762,304	(1) $p < .01$ (2) NS

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions: (1) column one versus column two, and (2) column one versus column three.

Table H.30

Major Course of Study
As a Function of Upward Bound Participation Pattern

Major Course Of Study	Upward Bound Participation Pattern[a/]			Significance Level Of Contrasts[b/]
	Typical	Atypical	Nonparticipant	
Education and Related Fields	(P = .274) P = .274	(P = .211) P = .205	(P = .125) P = .219	(1) NS (2) NS
Humanities	(P = .051) P = .051	(P = .117) P = .125	(P = .090) P = .101	(1) NS (2) NS
Other Nonscientific Academic Fields	(P = .343) P = .343	(P = .277) P = .269	(P = .327) P = .349	(1) NS (2) NS
Social Sciences	(P = .161) P = .161	(P = .220) P = .217	(P = .029) P = .037	(1) NS (2) $p < .01$
Natural Sciences, Engineering, and Related Fields	(P = .170) P = .170	(P = .176) P = .183	(P = .430) P = .294	(1) NS (2) NS
Number of Cases	304	81	102	
Weight Total	50,952	17,917	243,525	

NOTE: These analyses were restricted to PSE entrants in the 11th and 12th grade cohort who last attended a four-year institution. Two entries are provided in each cell: (1) raw weighted column proportions, P (in parentheses), and (2) adjusted weighted column proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group). Small cell sizes led to insufficient data for within stratum estimates in some strata; therefore, results should be viewed with caution.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.31

Rates of Reported Availability of Tutoring Services
At Institution Last Attended
As a Function of Upward Bound Participation Pattern
And Special Services Host Status of Institution

Special Services Host Status Of Institution Last Attended	Upward Bound Participation Pattern[a/]			Significance Level Of Contrasts[b/]
	Typical	Atypical	Nonparticipant	
Not Special Services Host	($\tilde{P}$ = .924) P = .924 N = 467 WN = 35,193	($\tilde{P}$ = .885) P = .881 N = 268 WN = 23,135	($\tilde{P}$ = .840) P = .794 N = 395 WN = 398,064	(1) NS (2) $p < .01$
Special Services Host	($\tilde{P}$ = .927) P = .927 N = 584 WN = 41,767	($\tilde{P}$ = .934) P = .925 N = 197 WN = 15,330	($\tilde{P}$ = .932) P = .931 N = 288 WN = 362,030	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, $\tilde{P}$ (in parentheses), (2) adjusted weighted proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.32

Rates of Reported Availability of Counseling Services
At Postsecondary Institutions Last Attended
As a Function of Upward Bound Participation Pattern
And Special Services Host Status of Institution

Special Services Host Status Of Institution Last Attended	Upward Bound Participation Pattern[a/] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b/]
Not Special Services Host	(P = .977) $\tilde{P}$ = .977 N = 471 WN = 35,234	(P = .987) $\tilde{P}$ = .999 N = 269 WN = 23,135	(P = .936) $\tilde{P}$ = .951 N = 395 WN = 398,064	(1) NS (2) NS
Special Services Host	(P = .985) $\tilde{P}$ = .985 N = 586 WN = 41,726	(P = .992) $\tilde{P}$ = .990 N = 197 WN = 15,330	(P = .995) $\tilde{P}$ = .980 N = 288 WN = 362,030	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

[a/] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[b/] Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.33

Rates of Reported Availability of Remedial Courses
At Postsecondary Institutions Last Attended
As a Function of Upward Bound Participation Pattern
And Special Services Host Status of Institution

Special Services Host Status Of Institution Last Attended	Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
	Typical	Atypical	Nonparticipant	
Not Special Services Host	(P = .947) $\tilde{P}$ = .947 N = 470 WN = 35,260	(P = .887) $\tilde{P}$ = .895 N = 269 WN = 23,232	(P = .869) $\tilde{P}$ = .894 N = 396 WN = 398,064	(1) $p < .05$ (2) NS
Special Services Host	(P = .964) $\tilde{P}$ = .964 N = 585 WN = 41,699	(P = .972) $\tilde{P}$ = .970 N = 195 WN = 15,233	(P = .962) $\tilde{P}$ = .965 N = 288 WN = 362,030	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.34

Reported Participation in Available Tutoring Services
As a Function of Upward Bound Participation Pattern
And Special Services Host Status of Institution Last Attended

Special Services Host Status	Participation in Tutoring Services	Upward Bound Participation Pattern[a/]			Significance Level Of Contrasts[b/]
		Typical	Atypical	Nonparticipant	
Not Host	Did Not Participate	(P = .326) P = .326	(P = .324) P = .327	(P = .520) P = .421	(1) NS (2) p < .05
	Participated Infrequently	(P = .244) P = .244	(P = .262) P = .235	(P = .159) P = .108	(1) NS (2) p < .01
	Participated Occasionally	(P = .286) P = .286	(P = .263) P = .299	(P = .262) P = .385	(1) NS (2) NS
	Participated Frequently	(P = .144) P = .144	(P = .151) P = .140	(P = .058) P = .087	(1) NS (2) NS
	Mean Value[c/]	(M = 2.25) M = 2.25	(M = 2.24) M = 2.25	(M = 1.86) M = 2.14	(1) NS (2) NS
	Number of Cases	427	232	323	
	Weight Total	32,524	20,475	334,557	
Host	Did Not Participate	(P = .294) P = .294	(P = .414) P = .402	(P = .525) P = .467	(1) p < .05 (2) p < .05
	Participated Infrequently	(P = .245) P = .245	(P = .205) P = .202	(P = .242) P = .234	(1) NS (2) NS
	Participated Occasionally	(P = .292) P = .292	(P = .241) P = .247	(P = .188) P = .212	(1) NS (2) NS
	Participated Frequently	(P = .169) P = .169	(P = .140) P = .150	(P = .046) P = .086	(1) NS (2) p < .05
	Mean Value[c/]	(M = 2.34) M = 2.34	(M = 2.11) M = 2.14	(M = 1.75) M = 1.92	(1) NS (2) p < .01
	Number of Cases	546	183	265	
	Weight Total	38,735	14,320	337,517	

NOTE: These analyses were restricted to PSE entrants who reported availability of tutoring services at institution last attended. The participation variable reported was computed as an aggregate variable over responses to participation in all specified tutoring services (see Appendix E). Two entries are provided in each cell: (1) raw weighted column proportions or means within major host-status breakdown, P or M (in parentheses), and (2) adjusted weighted column proportions or means, P or M (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions or means within each row: (1) column one versus column two, and (2) column one versus column three.

c/ Averages were taken for a four-valued scaling (from 1 to 4 in the order given) of extent of participation.

Table H.35

Reported Participation in Available Counseling Services
As a Function of Upward Bound Participation Pattern
And Special Services Host Status of Institution Last Attended

Special Services Host Status	Participation in Counseling Services	Upward Bound Participation Pattern[a/] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b/]
Not Host	Did Not Participate	(P = .249) P = .249	(P = .235) P = .245	(P = .374) P = .297	(1) NS (2) NS
	Participated Infrequently	(P = .421) P = .421	(P = .380) P = .360	(P = .432) P = .436	(1) NS (2) NS
	Participated Occasionally	(P = .274) P = .274	(P = .285) P = .292	(P = .168) P = .239	(1) NS (2) NS
	Participated Frequently	(P = .055) P = .055	(P = .101) P = .103	(P = .027) P = .028	(1) NS (2) NS
	Mean Value[c/]	(M = 2.14) M = 2.14	(M = 2.25) M = 2.25	(M = 1.85) M = 2.00	(1) NS (2) NS
	Number of Cases	461	265	377	
	Weight Total	34,415	22,831	372,467	
Host	Did Not Participate	(P = .199) P = .199	(P = .299) P = .259	(P = .500) P = .345	(1) $p < .05$ (2) $p < .05$
	Participated Infrequently	(P = .413) P = .413	(P = .395) P = .390	(P = .302) P = .294	(1) NS (2) $p < .05$
	Participated Occasionally	(P = .335) P = .335	(P = .233) P = .274	(P = .187) P = .324	(1) NS (2) NS
	Participated Frequently	(P = .052) P = .052	(P = .072) P = .077	(P = .011) P = .038	(1) NS (2) NS
	Mean Value[c/]	(M = 2.24) M = 2.24	(M = 2.08) M = 2.17	(M = 1.71) M = 2.05	(1) NS (2) NS
	Number of Cases	576	196	286	
	Weight Total	41,111	15,214	360,248	

NOTE: These analyses were restricted to PSE entrants who reported availability of counseling services at institution last attended. The participation variable reported was computed as an aggregate variable over responses to participation in all specified counseling services (see Appendix E). Two entries are provided in each cell: (1) raw weighted column proportions or means within major host-status breakdown, P or M (in parentheses), and (2) adjusted weighted column proportions or means, P or M (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

[a/] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[b/] Two *a priori* contrasts were formed from the adjusted column proportions or means within each row: (1) column one versus column two, and (2) column one versus column three.

[c/] Averages were taken for a four-valued scaling (from 1 to 4 in the order given) of extent of participation.

Table H.36

Reported Participation in Available Remedial Services
As a Function of Upward Bound Participation Pattern
And Special Services Host Status of Institution Last Attended

Special Services Host Status	Participation in Remedial Services	Upward Bound Participation Pattern[a/] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b/]
Not Host	Did Not Participate	(P = .384) P = .384	(P = .321) P = .311	(P = .526) P = .312	(1) NS (2) NS
	Participated Infrequently	(P = .286) P = .286	(P = .286) P = .305	(P = .271) P = .363	(1) NS (2) NS
	Participated Occasionally	(P = .260) P = .260	(P = .292) P = .284	(P = .134) P = .222	(1) NS (2) NS
	Participated Frequently	(P = .070) P = .070	(P = .101) P = .100	(P = .069) P = .102	(1) NS (2) NS
	Mean Value[c/]	(M = 2.02) M = 2.02	(M = 2.17) M = 2.17	(M = 1.75) M = 2.12	(1) NS (2) NS
	Number of Cases	444	241	351	
	Weight Total	33,396	20,599	346,067	
Host	Did Not Participate	(P = .363) P = .363	(P = .470) P = .426	(P = .648) P = .436	(1) NS (2) NS
	Participated Infrequently	(P = .307) P = .307	(P = .279) P = .298	(P = .211) P = .289	(1) NS (2) NS
	Participated Occasionally	(P = .243) P = .243	(P = .170) P = .189	(P = .092) P = .128	(1) NS (2) $p < .05$
	Participated Frequently	(P = .087) P = .087	(P = .081) P = .087	(P = .050) P = .146	(1) NS (2) NS
	Mean Value[c/]	(M = 2.05) M = 2.05	(M = 1.86) M = 1.94	(M = 1.54) M = 1.98	(1) NS (2) NS
	Number of Cases	562	188	275	
	Weight Total	40,188	14,807	348,280	

NOTE: These analyses were restricted to PSE entrants who reported availability of remedial services at institution last attended. The participation variable reported was computed as an aggregate variable over responses to participation in all specified remedial services (see Appendix E). Two entries are provided in each cell: (1) raw weighted column proportions or means within major host-status breakdown, P or M (in parentheses), and (2) adjusted weighted column proportions or means, P or M (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions or means within each row: (1) column one versus column two, and (2) column one versus column three.

c/ Averages were taken for a four-valued scaling (from 1 to 4 in the order given) of extent of participation.

Table H.37

Rates of Receiving Specified Types of Grants and Scholarships
During Most Recent Year of Postsecondary Education
As a Function of Upward Bound Participation Pattern

Type of Grant or Scholarship	Upward Bound Participation Pattern[a/]			Significance Level Of Contrasts[b/]
	Typical	Atypical	Nonparticipant	
Basic Educational Opportunity Grant	(P = .792) P = .792	(P = .667) P = .701	(P = .336) (P = .594)	(1) p < .01 (2) p < .01
Supplemental Educational Opportunity Grant	(P = .199) P = .199	(P = .144) P = .147	(P = .043) P = .078	(1) NS (2) p < .01
ROTC Scholarship Or Stipend	(P = .013) P = .013	(P = .006) P = .007	(P = .005) P = .006	(1) NS (2) NS
Other Federal Scholarship	(P = .018) P = .018	(P = .021) P = .021	(P = .008) P = .009	(1) NS (2) NS
Veterans Administration Programs	(P = .035) P = .035	(P = .035) P = .039	(P = .033) P = .045	(1) NS (2) NS
Social Security Benefits	(P = .115) P = .115	(P = .100) P = .096	(P = .089) P = .087	(1) NS (2) NS
State Scholarship	(P = .132) P = .132	(P = .075) P = .074	(P = .089) P = .091	(1) p < .01 (2) NS
Other Scholarship Or Grant	(P = .217) P = .217	(P = .161) P = .145	(P = .149) P = .097	(1) p < .05 (2) p < .01
Number of Cases	1,012	448	617	
Weight Total	76,960	38,467	762,304	

NOTE: These analyses were restricted to PSE entrants. Two entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), and (2) adjusted weighted proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.38

Rates of Receiving Specified Types of Loans
During Most Recent Year of Postsecondary Education
As a Function of Upward Bound Participation Pattern

Type of Loan	Upward Bound Participation Pattern[a] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b]
Federal Guaranteed Student Loan	(P = .083) $\tilde{P}$ = .083	(P = .087) $\tilde{P}$ = .088	(P = .080) $\tilde{P}$ = .056	(1) NS (2) NS
State Loan	(P = .020) $\tilde{P}$ = .020	(P = .014) $\tilde{P}$ = .016	(P = .010) $\tilde{P}$ = .020	(1) NS (2) NS
National Defense (Direct) Student Loan	(P = .272) $\tilde{P}$ = .272	(P = .220) $\tilde{P}$ = .228	(P = .062) $\tilde{P}$ = .104	(1) NS (2) $p < .01$
Other Loan	(P = .049) $\tilde{P}$ = .049	(P = .064) $\tilde{P}$ = .068	P = .058 $\tilde{P}$ = .071	(1) NS (2) NS
Number of Cases	1,007	446	614	
Weight Total	76,960	38,467	762,304	

NOTE: These analyses were restricted to PSE entrants. Two entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), and (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.39

Rates of Using Specified Types of Earnings or Savings
To Finance Most Recent Year of Postsecondary Education
As a Function of Upward Bound Participation Pattern

Type of Earning or Savings	Upward Bound Participation Pattern[a/]			Significance Level Of Contrasts[b/]
	Typical	Atypical	Nonparticipant	
Summer Earnings	($\tilde{P}$ = .213) P = .213	($\tilde{P}$ = .265) P = .250	($\tilde{P}$ = .320) P = .237	(1) NS (2) NS
College Sponsored Work Programs (While Taking Courses)	($\tilde{P}$ = .288) P = .288	($\tilde{P}$ = .239) P = .239	($\tilde{P}$ = .138) P = .161	(1) NS (2) $p < .01$
Other Earnings While Taking Courses	($\tilde{P}$ = .159) P = .159	($\tilde{P}$ = .182) P = .163	($\tilde{P}$ = .284) P = .176	(1) NS (2) NS
Number of Cases	970	425	592	
Weight Total	76,960	38,467	762,304	

NOTE: These analyses were restricted to PSE entrants. Two entries are provided in each cell: (1) raw weighted proportions, $\tilde{P}$ (in parentheses), and (2) adjusted weighted proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical-pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.40

Rates of Receiving Some Form of Grant During Most Recent Year of Postsecondary Education
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a]	Upward Bound Participation Pattern[b]			Significance Level Of Contrasts[c]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d]	(P = .864) P̃ = .864 N = 70 WN = 6,043	(P = .694) P̃ = .639 N = 64 WN = 6,716	(P = .470) P̃ = .571 N = 111 WN = 116,527	(1) p < .05 (2) p < .01
Two Year Academic Program	(P = .749) P̃ = .749 N = 190 WN = 14,544	(P = .710) P̃ = .741 N = 128 WN = 10,928	(P = .468) P̃ = .614 N = 184 WN = 225,805	(1) NS (2) p < .05
Four Year or Greater Academic Program	(P = .940) P̃ = .940 N = 758 WN = 56,373	(P = .856) P̃ = .892 N = 255 WN = 20,823	(P = .605) P̃ = .794 N = 325 WN = 419,972	(1) NS (2) p < .01

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, P̃ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

[a] Excluding graduate study.

[b] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[c] Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

[d] Small cell sizes in this row led to insufficient data for within stratum estimates in some strata; therefore, results should be viewed with caution.

Table H.41

Rates of Receiving Some Form of Loan During Most Recent Year of Postsecondary Education
As a Function of Upward Bound Participation Pattern
And Level of Offering of School Last Attended

Level of Offering of Institution Last Attended[a/]	Upward Bound Participation Pattern[b/]			Significance Level Of Contrasts[c/]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d/]	(P = .263) $\tilde{P}$ = .263 N = 69 WN = 5,943	(P = .382) $\tilde{P}$ = .412 N = 64 WN = 6,716	(P = .341) $\tilde{P}$ = .412 N = 110 WN = 116,002	(1) NS (2) NS
Two Year Academic Program	(P = .122) $\tilde{P}$ = .122 N = 190 WN = 14,568	(P = .145) $\tilde{P}$ = .158 N = 129 WN = 10,774	(P = .072) $\tilde{P}$ = .101 N = 182 WN = 225,329	(1) NS (2) NS
Four Year or Greater Academic Program	(P = .490) $\tilde{P}$ = .490 N = 758 WN = 56,449	(P = .476) $\tilde{P}$ = .483 N = 255 WN = 20,976	(P = .278) $\tilde{P}$ = .337 N = 325 WN = 420,973	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Excluding graduate study.

b/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

c/ Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

d/ Small cell sizes in this row led to insufficient data for within stratum estimates in some strata; therefore, results should be viewed with caution.

Table H.42

Rates of Using Some Form of Earnings or Savings
To Finance Most Recent Year of Postsecondary Education
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a]	Upward Bound Participation Pattern[b]			Significance Level Of Contrasts[c]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d]	(P = .505) $\tilde{P}$ = .505 N = 70 WN = 6,150	(P = .423) $\tilde{P}$ = .491 N = 64 WN = 6,716	(P = .376) $\tilde{P}$ = .370 N = 111 WN = 117,249	(1) NS (2) NS
Two Year Academic Program	(P = .486) $\tilde{P}$ = .486 N = 191 WN = 14,619	(P = .522) $\tilde{P}$ = .492 N = 129 WN = 10,774	(P = .539) $\tilde{P}$ = .421 N = 184 WN = 225,805	(1) NS (2) NS
Four Year or Greater Academic Program	(P = .596) $\tilde{P}$ = .596 N = 759 WN = 56,191	(P = .664) $\tilde{P}$ = .631 N = 256 WN = 20,976	(P = .721) $\tilde{P}$ = .615 N = 323 WN = 419,250	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

[a] Excluding graduate study.

[b] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[c] Two a priori contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

[d] Small cell sizes in this row led to insufficient data for within stratum estimates in some strata; therefore, results should be viewed with caution.

Terms of Postsecondary Education[a]	Aid Packaging	Participation P: Typical	Participation P: Atypical	Participation P: Nonparticipant	Significance Level Of Contrasts[c]
Eight or Fewer Terms	No Aid Received	(P = .033) P = .033	(P = .083) P = .084	(P = .147) P = .126	(1) p < .05 (2) p < .01
	Savings or Earnings Only	(P = .069) P = .069	(P = .155) P = .135	(P = .249) P = .167	(1) p < .05 (2) p < .01
	Loans Only	(P = .009) P = .009	(P = .010) P = .015	(P = .034) P = .042	(1) NS (2) NS
	Loans and Savings Or Earnings	(P = .026) P = .026	(P = .033) P = .028	(P = .035) P = .012	(1) NS (2) NS
	Grants or Scholarships Only	(P = .390) P = .390	(P = .243) P = .260	(P = .219) P = .335	(1) p < .01 (2) NS
	Grants or Scholarships And Savings or Earnings	(P = .213) P = .213	(P = .230) P = .222	(P = .198) P = .150	(1) NS (2) NS
	Grants or Scholarships And Loans	(P = .098) P = .098	(P = .131) P = .138	(P = .063) P = .094	(1) NS (2) NS
	Grants or Scholarships And Loans and Savings or Earnings	(P = .163) P = .163	(P = .113) P = .118	(P = .054) P = .076	(1) NS (2) p < .05
	Number of Cases	429	239	386	
	Weight Total	33,983	20,859	490,117	
Nine or More Terms	No Aid Received	(P = .016) P = .016	(P = .018) P = .015	(P = .057) P = .048	(1) NS (2) NS
	Savings or Earnings Only	(P = .048) P = .048	(P = .074) P = .044	(P = .305) P = .091	(1) NS (2) NS
	Loans Only	(P = .007) P = .007	(P = .010) P = .010	(P = .038) P = .015	(1) NS (2) NS
	Loans and Savings Or Earnings	(P = .007) P = .007	(P = .032) P = .034	(P = .038) P = .044	(1) NS (2) NS
	Grants or Scholarships Only	(P = .205) P = .205	(P = .226) P = .241	(P = .114) P = .218	(1) NS (2) NS
	Grants or Scholarships And Savings or Earnings	(P = .243) P = .243	(P = .225) P = .234	(P = .224) P = .260	(1) NS (2) NS
	Grants or Scholarships And Loans	(P = .127) P = .127	(P = .111) P = .121	(P = .037) P = .077	(1) NS (2) NS
	Grants or Scholarships And Loans and Savings or Earnings	(P = .350) P = .350	(P = .304) P = .303	(P = .188) P = .247	(1) NS (2) NS
	Number of Cases	583	207	229	
	Weight Total	33,983	20,859	490,117	
Nine or More Terms	No Aid Received	(P = .016) P = .016	(P = .018) P = .015	(P = .057) P = .048	(1) NS (2) NS
	Savings or Earnings Only	(P = .048) P = .048	(P = .074) P = .044	(P = .305) P = .091	(1) NS (2) NS
	Loans Only	(P = .007) P = .007	(P = .010) P = .010	(P = .038) P = .015	(1) NS (2) NS
	Loans and Savings Or Earnings	(P = .007) P = .007	(P = .032) P = .034	(P = .038) P = .044	(1) NS (2) NS
	Grants or Scholarships Only	(P = .205) P = .205	(P = .226) P = .241	(P = .114) P = .218	(1) NS (2) NS
	Grants or Scholarships And Savings or Earnings	(P = .243) P = .243	(P = .225) P = .234	(P = .224) P = .260	(1) NS (2) NS
	Grants or Scholarships And Loans	(P = .127) P = .127	(P = .111) P = .121	(P = .037) P = .077	(1) NS (2) NS
	Grants or Scholarships And Loans and Savings or Earnings	(P = .350) P = .350	(P = .304) P = .303	(P = .188) P = [illegible]	(1) NS (2) NS
	Number of Cases	583	207	229	
	Weight Total	42,928	17,607	272,187	

NOTE: These analyses were restricted to PSE entrants. Two entries are provided in each cell: (1) raw weighted column proportions within major terms-attended breakdown, P (in parentheses), and (2) adjusted weighted column proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

[a] Number of terms is quarter-system based; summer term attendance is included.

[b] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[c] Two *a priori* contrasts were formed from the adjusted column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.44

Average Grant Amount Received During Most Recent Year of Postsecondary Education
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a]	Upward Bound Participation Pattern[b]			Significance Level Of Contrasts[c]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d]	(M = $1,311) M = $1,311 N = 63 WN = 6,279	(M = $ 810) M = $ 843 N = 56 WN = 6,526	(M = $ 324) M = $ 522 N = 104 WN = 116,659	(1) NS (2) $p < .01$
Two Year Academic Program	(M = $ 747) M = $ 747 N = 169 WN = 14,806	(M = $ 618) M = $ 644 N = 115 WN = 11,116	(M = $ 316) M = $ 502 N = 173 WN = 223,978	(1) NS (2) $p < .01$
Four Year or Greater Academic Program	(M = $1,623) M = $1,623 N = 660 WN = 55,580	(M = $1,348 M = $1,360 N = 219 WN = 20,825	(M = $ 834) M = $1,088 N = 305 WN = 421,962	(1) $p < .05$ (2) $p < .01$

NOTE: These analyses were restricted to PSE entrants; the data were trimmed by excluding cases with suspect values (i.e., values greater than $8,000). Four entries are provided in each cell: (1) raw weighted means, M (in parentheses), (2) adjusted weighted means, M (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

[a] Excluding graduate study.

[b] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[c] Two a priori contrasts were formed from the adjusted means within each row: (1) column one versus column two, and (2) column one versus column three.

[d] Cell sizes in this row are sufficiently small to ensure that within stratum assumptions of the analytic model are violated; therefore, results should be viewed with extreme caution.

Table H.45

Average Loan Amount Received During Most Recent Year of Postsecondary Education
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a]	Upward Bound Participation Pattern[b]: Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[c]
Two Year or Less Vocational, Trade, or Business School[d]	(M = $272) M = $272 N = 67 WN = 5,801	(M = $468) M = $558 N = 64 WN = 7,087	(M = $447) M = $494 N = 109 WN = 117,574	(1) NS (2) p < .05
Two Year Academic Program	(M = $ 93) M = $ 93 N = 188 WN = 14,702	(M = $ 96) M = $ 99 N = 128 WN = 11,146	(M = $ 47) M = $ 55 N = 181 WN = 225,559	(1) NS (2) NS
Four Year or Greater Academic Program	(M = $486) M = $486 N = 723 WN = 56,457	(M = $387 M = $403 N = 239 WN = 20,234	(M = $303) M = $376 N = 322 WN = 419,171	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants; the data were trimmed by excluding cases with suspect values (i.e., values greater than $4,800). Four entries are provided in each cell: (1) raw weighted means, M (in parentheses), (2) adjusted weighted means, M (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

[a] Excluding graduate study.

[b] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[c] Two a priori contrasts were formed from the adjusted means within each row: (1) column one versus column two, and (2) column one versus column three.

[d] Small cell sizes in this row led to insufficient data for within cell estimates in some strata; therefore, results should be viewed with caution.

Table H.46

Average Amount of Earnings or Savings
Used to Finance Most Recent Year of Postsecondary Education
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a/]	Upward Bound Participation Pattern[b/]			Significance Level Of Contrasts[c/]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d/]	($\underset{\sim}{M}$ = \$282) $\tilde{M}$ = \$282 N = 62 WN = 5,846	($\underset{\sim}{M}$ = \$174) $\tilde{M}$ = \$211 N = 58 WN = 6,645	($\underset{\sim}{M}$ = \$270) $\tilde{M}$ = \$239 N = 106 WN = 123,641	(1) NS (2) NS
Two Year Academic Program	($\underset{\sim}{M}$ = \$210) $\tilde{M}$ = \$210 N = 172 WN = 14,764	($\underset{\sim}{M}$ = \$378) $\tilde{M}$ = \$360 N = 119 WN = 10,507	($\underset{\sim}{M}$ = \$323) $\tilde{M}$ = \$209 N = 165 WN = 228,474	(1) $p < .05$ (2) NS
Four Year or Greater Academic Program	($\underset{\sim}{M}$ = \$442) $\tilde{M}$ = \$442 N = 692 WN = 56,349	($\underset{\sim}{M}$ = \$492) $\tilde{M}$ = \$459 N = 236 WN = 21,316	($\underset{\sim}{M}$ = \$660) $\tilde{M}$ = \$552 N = 302 WN = 410,189	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants; the data were trimmed by excluding cases with suspect values (i.e., values greater than \$7,000). Four entries are provided in each cell: (1) raw weighted means, $\underset{\sim}{M}$ (in parentheses), (2) adjusted weighted means, $\tilde{M}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Excluding graduate study.

b/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

c/ Two a priori contrasts were formed from the adjusted means within each row: (1) column one versus column two, and (2) column one versus column three.

d/ Small cell sizes in this row led to insufficient data for within stratum estimates in some strata; therefore, results should be viewed with caution.

Table H.47

Reported Adequacy of Financial Assistance During Last Year of Postsecondary Study
From Scholarships and Grants
As a Function of Upward Bound Participation Pattern
And Cost Category of Institution Last Attended

Cost Category	Adequacy of Scholarships and Grants	Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
		Typical	Atypical	Nonparticipant	
Less than $1,000	No Support Needed	(P = .100) P = .100	(P = .108) P = .067	(P = .361) P = .178	(1) NS (2) p < .05
	Very Adequate	(P = .257) P = .257	(P = .228) P = .239	(P = .122) P = .183	(1) NS (2) NS
	Adequate	(P = .366) P = .366	(P = .338) P = .348	(P = .201) P = .291	(1) NS (2) NS
	Inadequate	(P = .142) P = .142	(P = .097) P = .099	(P = .079) P = .100	(1) NS (2) NS
	Not Received Or Very Inadequate	(P = .135) P = .135	(P = .228) P = .247	(P = .237) P = .247	(1) p < .01 (2) p < .05
	Number of Cases	726	331	528	
	Weight Total	54,552	27,962	637,368	
$1,000 Or More	No Support Needed	(P = .044) P = .044	(P = .022) P = .000	(P = .243) P = .030	(1) NS (2) NS
	Very Adequate	(P = .335) P = .335	(P = .446) P = .432	(P = .076) P = .126	(1) NS (2) p < .05
	Adequate	(P = .418) P = .418	(P = .304) P = .339	(P = .229) P = .349	(1) NS (2) NS
	Inadequate	(P = .126) P = .126	(P = .130) P = .131	(P = .216) P = .296	(1) NS (2) p < .05
	Not Received Or Very Inadequate	(P = .077) P = .077	(P = .098) P = .099	(P = .236) P = .199	(1) NS (2) p < .05
	Number of Cases	302	109	107	
	Weight Total	22,408	10,606	124,936	

NOTE: These analyses were restricted to PSE entrants. Two entries are provided in each cell: (1) raw weighted column proportions within major cost category breakdown, P (in parentheses), and (2) adjusted weighted column proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.48

Rates of Reporting Grants and Scholarships
As Unneeded or Adequate
During Last Year of Postsecondary Study
As a Function of Upward Bound Participation Pattern
And Cost Category of Institution Last Attended

Cost Category of Institution Last Attended	Upward Bound Participation Pattern[a/]			Significance Level Of Contrasts[b/]
	Typical	Atypical	Nonparticipant	
Less than $1,000	(P = .724) P̃ = .724 N = 726 WN = 54,552	(P = .674) P̃ = .654 N = 331 WN = 27,962	(P = .685) P̃ = .652 N = 528 WN = 637,368	(1) NS (2) NS
$1,000 or More	(P = .797) P̃ = .797 N = 302 WN = 22,408	(P = .772) P̃ = .770 N = 109 WN = 10,606	(P = .548) P̃ = .505 N = 107 WN = 124,936	(1) NS (2) $p < .01$

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, P̃ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two a priori contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.49

Reported Adequacy of Financial Assistance During Last Year of Postsecondary Study From Loans
As a Function of Upward Bound Participation Pattern
And Cost Category of Institution Last Attended

Cost Category	Adequacy of Loans	Upward Bound Participation Pattern[a] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b]
Less than $1,000	No Support Needed	(P = .315) P = .315	(P = .311) P = .277	(P = .541) P = .393	(1) NS (2) NS
	Very Adequate	(P = .095) P = .095	(P = .071) P = .074	(P = .039) P = .051	(1) NS (2) $p < .05$
	Adequate	(P = .197) P = .197	(P = .162) P = .171	(P = .112) P = .127	(1) NS (2) $p < .05$
	Inadequate	(P = .062) P = .062	(P = .061) P = .063	(P = .027) P = .033	(1) NS (2) NS
	Not Received Or Very Inadequate	(P = .331) P = .331	(P = .394) P = .414	(P = .281) P = .396	(1) $p < .05$ (2) NS
	Number of Cases	725	331	530	
	Weight Total	54,552	27,886	637,368	
$1,000 Or More	No Support Needed	(P = .209) P = .209	(P = .188) P = .143	(P = .396) P = .243	(1) NS (2) NS
	Very Adequate	(P = .147) P = .147	(P = .331) P = .331	(P = .078) P = .167	(1) NS (2) NS
	Adequate	(P = .329) P = .329	(P = .216) P = .224	(P = .241) P = .240	(1) NS (2) NS
	Inadequate	(P = .084) P = .084	(P = .052) P = .050	(P = .050) P = .043	(1) NS (2) NS
	Not Received Or Very Inadequate	(P = .231) P = .231	(P = .213) P = .251	(P = .234) P = .307	(1) NS (2) NS
	Number of Cases	301	110	107	
	Weight Total	22,408	10,682	124,936	

NOTE: These analyses were restricted to PSE entrants. Two entries are provided in each cell: (1) raw weighted column proportions within major cost category breakdown, P (in parentheses), and (2) adjusted weighted column proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age for grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.50

Rates of Reporting Loans
As Unneeded or Adequate
During Last Year of Postsecondary Study
As a Function of Upward Bound Participation Pattern
And Cost Category of Institution Last Attended

Cost Category of Institution Last Attended	Upward Bound Participation Pattern[a/]			Significance Level Of Contrasts[b/]
	Typical	Atypical	Nonparticipant	
Less than $1,000	($\tilde{P}$ = .607) $\tilde{P}$ = .607 N = 725 WN = 54,552	($\tilde{P}$ = .545) $\tilde{P}$ = .523 N = 331 WN = 27,886	($\tilde{P}$ = .692) $\tilde{P}$ = .572 N = 530 WN = 637,368	(1) $p < .05$ (2) NS
$1,000 or More	($\tilde{P}$ = .685) $\tilde{P}$ = .685 N = 301 WN = 22,408	($\tilde{P}$ = .735) $\tilde{P}$ = .698 N = 110 WN = 10,682	($\tilde{P}$ = .716) $\tilde{P}$ = .650 N = 107 WN = 124,936	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

[a/] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[b/] Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.51

Reported Adequacy of Financial Assistance During Last Year of Postsecondary Study From Work Opportunities
As a Function of Upward Bound Participation Pattern
And Cost Category of Institution Last Attended

Cost Category	Adequacy of Work Opportunities	Upward Bound Participation Pattern[a/] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b/]
Less than $1,000	No Support Needed	(P = .181) P = .181	(P = .195) P = .168	(P = .281) P = .183	(1) NS (2) NS
	Very Adequate	(P = .123) P = .123	(P = .113) P = .102	(P = .125) P = .063	(1) NS (2) NS
	Adequate	(P = .235) P = .235	(P = .280) P = .281	(P = .241) P = .222	(1) NS (2) NS
	Inadequate	(P = .155) P = .155	(P = .150) P = .159	(P = .136) P = .187	(1) NS (2) NS
	Not Received Or Very Inadequate	(P = .306) P = .306	(P = .262) P = .290	(P = .217) P = .345	(1) NS (2) NS
	Number of Cases	724	330	529	
	Weight Total	54,641	27,886	637,368	
$1,000 Or More	No Support Needed	(P = .111) P = .111	(P = .100) P = .106	(P = .223) P = .136	(1) NS (2) NS
	Very Adequate	(P = .141) P = .141	(P = .138) P = .126	(P = .151) P = .102	(1) NS (2) NS
	Adequate	(P = .312) P = .312	(P = .444) P = .420	(P = .221) P = .228)	(1) NS (2) NS
	Inadequate	(P = .201) P = .201	(P = .139) P = .121	(P = .200) P = .179	(1) NS (2) NS
	Not Received Or Very Inadequate	(P = .235) P = .235	(P = .180) P = .226	(P = .205) P = .355	(1) NS (2) NS
	Number of Cases	300	110	107	
	Weight Total	22,319	10,682	124,936	

NOTE: These analyses were restricted to PSE entrants. Two entries are provided in each cell: (1) raw weighted column proportions within major cost category breakdown, P (in parentheses), and (2) adjusted weighted column proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation patte[illegible]s defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.52

Rates of Reporting Work Opportunities
As Unneeded or Adequate
During Last Year of Postsecondary Study
As a Function of Upward Bound Participation Pattern
And Cost Category of Institution Last Attended

Cost Category of Institution Last Attended	Upward Bound Participation Pattern[a/]			Significance Level Of Contrasts[b/]
	Typical	Atypical	Nonparticipant	
Less than $1,000	($\tilde{P}$ = .539) $\tilde{P}$ = .539 N = 724 WN = 54,641	($\tilde{P}$ = .588) $\tilde{P}$ = .551 N = 330 WN = 27,886	($\tilde{P}$ = .648) $\tilde{P}$ = .468 N = 529 WN = 637,368	(1) NS (2) NS
$1,000 or More	($\tilde{P}$ = .563) $\tilde{P}$ = .563 N = 300 WN = 22,319	($\tilde{P}$ = .681) $\tilde{P}$ = .653 N = 110 WN = 10,682	($\tilde{P}$ = .595) $\tilde{P}$ = .466 N = 107 WN = 124,936	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.53

Reported Adequacy of Financial Assistance During Last Year of Postsecondary Study
From Total of Grants, Scholarships, Loans, and Work Opportunities
As a Function of Upward Bound Participation Pattern
And Cost Category of Institution Last Attended

Cost Category	Adequacy of Grants, Scholarships, Loans, and Work Opportunities	Upward Bound Participation Pattern[a/]			Significance Level Of Contrasts[b/]
		Typical	Atypical	Nonparticipant	
Less than $1,000	No Support Needed	(P = .082) P = .082	(P = .098) P = .063	(P = .280) P = .129	(1) NS (2) NS
	Very Adequate	(P = .173) P = .173	(P = .135) P = .134	(P = .085) P = .068	(1) NS (2) NS
	Adequate	(P = .387) P = .387	(P = .405) P = .420	(P = .322) P = .367	(1) NS (2) NS
	Inadequate	(P = .235) P = .235	(P = .212) P = .215	(P = .171) P = .220	(1) NS (2) NS
	Not Received Or Very Inadequate	(P = .123) P = .123	(P = .150) P = .168	(P = .142) P = .216	(1) NS (2) $p < .01$
	Number of Cases	751	346	564	
	Weight Total	54,512	28,272	641,330	
$1,000 Or More	No Support Needed	(P = .057) P = .057	(P = .020) P = .023	(P = .165) P = .054	(1) NS (2) NS
	Very Adequate	(P = .188) P = .188	(P = .394) P = .394	(P = .137) P = .163	(1) NS (2) NS
	Adequate	(P = .501) P = .501	(P = .367) P = .379	(P = .314) P = .369	(1) NS (2) NS
	Inadequate	(P = .202) P = .202	(P = .195) P = .160	(P = .267) P = .214	(1) NS (2) NS
	Not Received Or Very Inadequate	(P = .053) P = .053	(P = .024) P = .045	(P = .116) P = .200	(1) NS (2) $p < .05$
	Number of Cases	306	110	109	
	Weight Total	22,448	10,295	120,974	

NOTE: These analyses were restricted to PSE entrants. Two entries are provided in each cell: (1) raw weighted column proportions within major cost category breakdown, P (in parentheses), and (2) adjusted weighted column proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.54

Rates of Reporting Total of Grants, Scholarships, Loans, and Work Opportunities
As Unneeded or Adequate
During Last Year of Postsecondary Study
As a Function of Upward Bound Participation Pattern
And Cost Category of Institution Last Attended

Cost Category of Institution Last Attended	Upward Bound Participation Pattern[a/]			Significance Level Of Contrasts[b/]
	Typical	Atypical	Nonparticipant	
Less than $1,000	(P = .642) $\tilde{P}$ = .642 N = 751 WN = 54,512	(P = .638) $\tilde{P}$ = .617 N = 346 WN = 28,272	(P = .687) $\tilde{P}$ = .564 N = 564 WN = 641,330	(1) NS (2) NS
$1,000 or More	(P = .745) $\tilde{P}$ = .745 N = 306 WN = 22,448	(P = .781) $\tilde{P}$ = .795 N = 110 WN = 10,295	(P = .617) $\tilde{P}$ = .586 N = 109 WN = 120,974	(1) NS (2) $p < .05$

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

[a/] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[b/] Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.55

Reported Adequacy of Financial Assistance During Last Year of Postsecondary Study From Family Contributions
As a Function of Upward Bound Participation Pattern
And Cost Category of Institution Last Attended

Cost Category	Adequacy of Family Contributions	Upward Bound Participation Pattern[a]: Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[b]
Less than $1,000	No Support Needed	(P = .235) P = .235	(P = .182) P = .185	(P = .194) P = .192	(1) NS (2) NS
	Very Adequate	(P = .101) P = .101	(P = .123) P = .093	(P = .263) P = .140	(1) NS (2) NS
	Adequate	(P = .190) P = .190	(P = .208) P = .215	(P = .275) P = .302	(1) NS (2) p < .05
	Inadequate	(P = .133) P = .133	(P = .145) P = .147	(P = .068) P = .094	(1) NS (2) NS
	Not Received Or Very Inadequate	(P = .340) P = .340	(P = .342) P = .360	(P = .199) P = .272	(1) NS (2) NS
	Number of Cases	704	328	531	
	Weight Total	53,782	28,115	637,883	
$1,000 Or More	No Support Needed	(P = .164) P = .164	(P = .368) P = .344	(P = .158) P = .175	(1) NS (2) NS
	Very Adequate	(P = .080) P = .080	(P = .054) P = .088	(P = .214) P = .144	(1) NS (2) NS
	Adequate	(P = .204) P = .204	(P = .209) P = .216	(P = .338) P = .394	(1) NS (2) NS
	Inadequate	(P = .142) P = .142	(P = .052) P = .044	(P = .069) P = .041	(1) p < .05 (2) p < .05
	Not Received Or Very Inadequate	(P = .410) P = .410	(P = .317) P = .308	(P = .220) P = .245	(1) NS (2) p < .05
	Number of Cases	295	105	108	
	Weight Total	23,083	10,452	124,516	

NOTE: These analyses were restricted to PSE entrants. Two entries are provided in each cell: (1) raw weighted column proportions within major cost category breakdown, P (in parentheses), and (2) adjusted weighted column proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group).

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.56

Rates of Reporting Family Contributions
As Unneeded or Adequate
During Last Year of Postsecondary Study
As a Function of Upward Bound Participation Pattern
And Cost Category of Institution Last Attended

Cost Category of Institution Last Attended	Upward Bound Participation Pattern[a/]			Significance Level Of Contrasts[b/]
	Typical	Atypical	Nonparticipant	
Less than $1,000	($\tilde{P}$ = .526) P = .526 N = 704 WN = 53,782	($\tilde{P}$ = .513) P = .493 N = 328 WN = 28,115	($\tilde{P}$ = .733) P = .635 N = 531 WN = 637,883	(1) NS (2) $p < .05$
$1,000 or More	($\tilde{P}$ = .449) P = .449 N = 295 WN = 23,083	($\tilde{P}$ = .631) P = .649 N = 105 WN = 10,452	($\tilde{P}$ = .711) P = .713 N = 108 WN = 124,516	(1) NS (2) $p < .01$

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, $\tilde{P}$ (in parentheses), (2) adjusted weighted proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

[a/] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[b/] Two a priori contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Table H.57

Average Percent of Total Idealized Financial Aid Package That Would Come From Loans and Work Study As a Function of Upward Bound Participation Pattern

Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
Typical	Atypical	Nonparticipant	
(M = 28.7%) M = 28.7% N = 1,001 WN = 76,889	(M = 31.4%) M = 31.5% N = 424 WN = 38,465	(M = 26.3%) M = 27.7% N = 629 WN = 762,374	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted means, M (in parentheses), (2) adjusted weighted means, M (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted means: (1) column one versus column two, and (2) column one versus column three.

Table H.58

Average Number of Terms Attending Postsecondary Institution
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a]	Upward Bound Participation Pattern[b]			Significance Level Of Contrasts[c]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d]	(M = 5.78) $\tilde{M}$ = 5.78 N = 91 WN = 6,398	(M = 4.69) $\tilde{M}$ = 4.45 N = 78 WN = 6,965	(M = 4.86) $\tilde{M}$ = 4.32 N = 136 WN = 129,636	(1) $p < .05$ (2) NS
Two Year Academic Program	(M = 6.95) $\tilde{M}$ = 6.95 N = 215 WN = 14,717	(M = 6.25) $\tilde{M}$ = 6.31 N = 138 WN = 10,627	(M = 5.51) $\tilde{M}$ = 5.99 N = 222 WN = 206,091	(1) NS (2) NS
Four Year or Greater Academic Program	(M = 9.91) $\tilde{M}$ = 9.91 N = 856 WN = 55,792	(M = 9.62) $\tilde{M}$ = 9.74 N = 289 WN = 20,976	(M = 8.54) $\tilde{M}$ = 9.59 N = 376 WN = 424,367	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted means, M (in parentheses), (2) adjusted weighted means, $\tilde{M}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

[a] Excluding graduate study.

[b] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[c] Two *a priori* contrasts were formed from the adjusted means within each row: (1) column one versus column two, and (2) column one versus column three.

[d] Small cell sizes in this row led to insufficient data for within stratum estimates in some strata; therefore, results should be viewed with caution.

Table H.59

Proportions of Individuals Graduated From or Currently Enrolled in a Postsecondary Institution
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a]	Upward Bound Participation Pattern[b]			Significance Level Of Contrasts[c]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d]	(P = .705) $\tilde{P}$ = .705 N = 91 WN = 6,416	(P = .720) $\tilde{P}$ = .664 N = 78 WN = 7,029	(P = .692) $\tilde{P}$ = .597 N = 135 WN = 130,236	(1) NS (2) NS
Two Year Academic Program	(P = .519) $\tilde{P}$ = .519 N = 213 WN = 14,663	(P = .548) $\tilde{P}$ = .548 N = 134 WN = 10,657	(P = .537) $\tilde{P}$ = .496 N = 220 WN = 207,354	(1) NS (2) NS
Four Year or Greater Academic Program	(P = .663) $\tilde{P}$ = .663 N = 849 WN = 55,828	(P = .707) $\tilde{P}$ = .714 N = 286 WN = 20,882	(P = .811) $\tilde{P}$ = .781 N = 370 WN = 424,532	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Excluding graduate study.

b/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

c/ Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

d/ Small cell sizes in this row led to insufficient data for within stratum estimates in some strata; therefore, results should be viewed with caution.

Table H.60

Proportions of Individuals Graduated From, Currently Enrolled In, or Intending to Return To a Postsecondary Institution as a Function of Upward Bound Participation Pattern And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a]	Upward Bound Participation Pattern[b]			Significance Level Of Contrasts[c]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d]	(P = .914) P = .914 N = 91 WN = 6,398	(P = .823) P = .798 N = 78 WN = 7,167	(P = .870) P = .838 N = 136 WN = 131,767	(1) NS (2) NS
Two Year Academic Program	(P = .824) P = .824 N = 215 WN = 14,706	(P = .840) P = .844 N = 136 WN = 10,657	(P = .782) P = .810 N = 220 WN = 205,177	(1) NS (2) NS
Four Year or Greater Academic Program	(P = .884) P = .884 N = 852 WN = 55,802	(P = .885) P = .894 N = 285 WN = 20,744	(P = .928) P = .929 N = 370 WN = 425,177	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

[a] Excluding graduate study.

[b] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[c] Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

[d] Small cell sizes in this row led to insufficient data for within stratum estimates in some strata; therefore, results should be viewed with caution.

Table H.61

Rates of Full-Time-Only Attendance as a Function of Upward Bound Participation Pattern And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a/]	Upward Bound Participation Pattern[b/]			Significance Level Of Contrasts[c/]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d/]	($\tilde{P}$ = .781) P = .781 N = 93 WN = 6,525	($\tilde{P}$ = .731) P = .686 N = 78 WN = 6,965	($\tilde{P}$ = .721) P = .678 N = 135 WN = 128,204	(1) NS (2) NS
Two Year Academic Program	($\tilde{P}$ = .646) P = .646 N = 213 WN = 14,487	($\tilde{P}$ = .683) P = .677 N = 138 WN = 10,518	($\tilde{P}$ = .548) P = .545 N = 222 WN = 207,718	(1) NS (2) NS
Four Year or Greater Academic Program	($\tilde{P}$ = .884) P = .884 N = 855 WN = 55,895	($\tilde{P}$ = .833) P = .837 N = 289 WN = 21,085	($\tilde{P}$ = .846) P = .880 N = 373 WN = 426,200	(1) NS (2) NS

NOTE: These analyses were conducted for PSE entrants only. Four entries are provided in each cell: (1) raw weighted proportions, $\tilde{P}$ (in parentheses), (2) adjusted weighted proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Excluding graduate study.

b/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

c/ Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

d/ Small cell sizes in this row led to insufficient data for within stratum estimates in some strata; therefore, results should be viewed with caution.

Table H.62

Average Percent of Available Summer Term Sessions Attended
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a]	Upward Bound Participation Pattern[b]			Significance Level Of Contrasts[c]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d]	($\tilde{M}$ = 36.7%) $\tilde{M}$ = 36.7% N = 91 WN = 6,398	($\tilde{M}$ = 46.6%) $\tilde{M}$ = 49.9% N = 78 WN = 6,965	($\tilde{M}$ = 40.1%) $\tilde{M}$ = 46.3% N = 136 WN = 129,636	(1) NS (2) NS
Two Year Academic Program	($\tilde{M}$ = 27.7%) $\tilde{M}$ = 27.7% N = 215 WN = 14,717	($\tilde{M}$ = 23.4%) $\tilde{M}$ = 23.3% N = 138 WN = 10,627	($\tilde{M}$ = 20.7%) $\tilde{M}$ = 22.3% N = 222 WN = 206,091	(1) NS (2) NS
Four Year or Greater Academic Program	($\tilde{M}$ = 40.2%) $\tilde{M}$ = 40.2% N = 856 WN = 55,792	($\tilde{M}$ = 34.8%) $\tilde{M}$ = 35.5% N = 289 WN = 20,976	($\tilde{M}$ = 23.6%) $\tilde{M}$ = 30.7% N = 376 WN = 424,367	(1) NS (2) $p < .05$

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted means, $\tilde{M}$ (in parentheses), (2) adjusted weighted means, $\tilde{M}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

[a] Excluding graduate study.

[b] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[c] Two *a priori* contrasts were formed from the adjusted means within each row: (1) column one versus column two, and (2) column one versus column three.

[d] Small cell sizes in this row led to insufficient data for within stratum estimates in some strata; therefore, results should be viewed with caution.

Table H.63

Average Postsecondary Grade Point Average at Institution Last Attended
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a]	Upward Bound Participation Pattern[b]			Significance Level Of Contrasts[c]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d]	(M = 2.13) M = 2.13 N = 26 WN = 2,395	(M = 1.86) M = 1.81 N = 16 WN = 1,936	(M = 2.40) M = 2.63 N = 26 WN = 55,728	(1) NS (2) $p < .05$
Two Year Academic Program	(M = 1.91) M = 1.91 N = 141 WN = 13,641	(M = 1.69) M = 1.74 N = 89 WN = 11,726	(M = 2.07) M = 1.95 N = 150 WN = 208,798	(1) NS (2) NS
Four Year or Greater Academic Program	(M = 2.05) M = 2.05 N = 695 WN = 62,878	(M = 1.98) M = 1.92 N = 224 WN = 24,805	(M = 2.40) M = 2.19 N = 294 WN = 495,824	(1) NS (2) $p < .05$

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted means, M (in parentheses), (2) adjusted weighted means, M (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Excluding graduate study.

b/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

c/ Two a priori contrasts were formed from the adjusted means within each row: (1) column one versus column two, and (2) column one versus column three.

d/ Cell sizes in this row are sufficiently small to ensure that within stratum assumptions of the analytic model are violated; therefore, results should be viewed with extreme caution.

Table H.64

Average Postsecondary Grade Point Average for Non-Transferring Students
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Attended

Level of Offering of Institution Attended[a]	Upward Bound Participation Pattern[b]			Significance Level Of Contrasts[c]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d]	(M = 2.07) M = 2.07 N = 23 WN = 2,168	(M = 1.72) M = 1.66 N = 13 WN = 1,622	(M = 2.44) M = 2.63 N = 23 WN = 52,399	(1) NS (2) $p < .05$
Two Year Academic Program[e]	(M = 1.79) M = 1.79 N = 121 WN = 11,784	(M = 1.66) M = 1.71 N = 78 WN = 10,369	(M = 2.04) M = 1.91 N = 138 WN = 197,667	(1) NS (2) NS
Four Year or Greater Academic Program	(M = 1.94) M = 1.94 N = 135 WN = 46,552	(M = 1.97) M = 1.92 N = 306 WN = 19,421	(M = 2.29) M = 2.21 N = 451 WN = 331,106	(1) NS (2) $p < .01$

NOTE: These analyses were restricted to PSE entrants who did not transfer from the originally entered institution. Four entries are provided in each cell: (1) raw weighted means, M (in parentheses), (2) adjusted weighted means, M (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Excluding graduate study.

b/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

c/ Two a priori contrasts were formed from the adjusted means within each row: (1) column one versus column two, and (2) column one versus column three.

d/ Cell sizes in this row are sufficiently small to ensure that within stratum assumptions of the analytic model are violated; therefore, results should be viewed with extreme caution.

e/ Small cell sizes in this row led to insufficient data for within stratum estimates in some strata; therefore, results should be interpreted with caution.

Table H.65

Rates of Being Placed on Academic Probation
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a/]	Upward Bound Participation Pattern[b/]			Significance Level Of Contrasts[c/]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d/]	(P = .161) $\tilde{P}$ = .161 N = 59 WN = 5,275	(P = .000) $\tilde{P}$ = .060 N = 45 WN = 5,737	(P = .109) $\tilde{P}$ = .025 N = 71 WN = 95,136	(1) p < .01 (2) p < .01
Two Year Academic Program	(P = .178) $\tilde{P}$ = .178 N = 165 WN = 13,743	(P = .151) $\tilde{P}$ = .128 N = 100 WN = 10,913	(P = .149) $\tilde{P}$ = .135 N = 173 WN = 209,091	(1) NS (2) NS
Four Year or Greater Academic Program	(P = .247) $\tilde{P}$ = .247 N = 718 WN = 57,775	(P = .238) $\tilde{P}$ = .260 N = 236 WN = 21,817	(P = .188) $\tilde{P}$ = .219 N = 302 WN = 458,243	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Excluding graduate study.

b/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

c/ Two <u>a priori</u> contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

d/ Cell sizes in this row were sufficiently small to ensure that within stratum assumptions of the analytic model are violated; therefore, results should be viewed with extreme caution.

Table H.66

Rates of Academic Dismissal
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a]	Upward Bound Participation Pattern[b]			Significance Level Of Contrasts[c]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d]	($\tilde{P}$ = .000) P = .000 N = 59 WN = 5,275	($\tilde{P}$ = .051) P = .070 N = 45 WN = 5,737	($\tilde{P}$ = .107) P = .147 N = 71 WN = 95,136	(1) NS (2) NS
Two Year Academic Program	($\tilde{P}$ = .058) P = .058 N = 165 WN = 13,743	($\tilde{P}$ = .073) P = .074 N = 100 WN = 10,913	($\tilde{P}$ = .005) P = .010 N = 173 WN = 209,091	(1) NS (2) NS
Four Year or Greater Academic Program	($\tilde{P}$ = .095) P = .095 N = 718 WN = 57,775	($\tilde{P}$ = .079) P = .083 N = 236 WN = 21,817	($\tilde{P}$ = .066) P = .060 N = 302 WN = 458,243	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants. Four entries are provided in each cell: (1) raw weighted proportions, $\tilde{P}$ (in parentheses), (2) adjusted weighted proportions, P (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Excluding graduate study.

b/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

c/ Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

d/ Cell sizes in this row are sufficiently small to ensure that within stratum assumptions of the analytic model are violated; therefore, results should be viewed with extreme caution.

Table H.67

Average Number of Earned Postsecondary Credits
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a]	Upward Bound Participation Pattern[b] Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[c]
Two Year or Less Vocational, Trade, or Business School[d]	($\tilde{M}$ = 30.0) $\tilde{M}$ = 30.0 N = 29 WN = 2,586	($\tilde{M}$ = 39.4) $\tilde{M}$ = 38.1 N = 18 WN = 2,094	($\tilde{M}$ = 38.9) $\tilde{M}$ = 48.8 N = 28 WN = 58,760	(1) NS (2) NS
Two Year Academic Program	($\tilde{M}$ = 38.9) $\tilde{M}$ = 38.9 N = 163 WN = 14,247	($\tilde{M}$ = 36.7) $\tilde{M}$ = 39.5 N = 100 WN = 12,297	($\tilde{M}$ = 31.3) $\tilde{M}$ = 31.5 N = 172 WN = 223,065	(1) NS (2) NS
Four Year or Greater Academic Program	($\tilde{M}$ = 101.4) $\tilde{M}$ = 101.4 N = 700 WN = 58,664	($\tilde{M}$ = 102.5) $\tilde{M}$ = 101.9 N = 232 WN = 23,644	($\tilde{M}$ = 103.8) $\tilde{M}$ = 105.9 N = 296 WN = 464,650	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants; the data were trimmed by excluding cases with suspect values (i.e., more than 220 earned credits). Four entries are provided in each cell: (1) raw weighted means, $\tilde{M}$ (in parentheses), (2) adjusted weighted means, $\tilde{M}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), and (3) cell size, N, and (4) total cell weight, WN.

a/ Excluding graduate study.

b/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

c/ Two a priori contrasts were formed from the adjusted means within each row: (1) column one versus column two, and (2) column one versus column three.

d/ Cell sizes in this row are sufficiently small to ensure that within stratum assumptions of the analytic model are violated; therefore, results should be viewed with extreme caution.

Table H.68

Average Credits Earned Relative to Credits Expected Under Normal Progression
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a]	Upward Bound Participation Pattern[b]			Significance Level Of Contrasts[c]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d]	(M = .306) $\tilde{M}$ = .306 N = 29 WN = 2,586	(M = .331) $\tilde{M}$ = .307 N = 18 WN = 2,094	(M = .367) $\tilde{M}$ = .410 N = 29 WN = 59,564	(1) NS (2) NS
Two Year Academic Program	(M = .310) $\tilde{M}$ = .310 N = 163 WN = 14,247	(M = .273) $\tilde{M}$ = .291 N = 100 WN = 12,297	(M = .271) $\tilde{M}$ = .259 N = 171 WN = 221,839	(1) NS (2) NS
Four Year or Greater Academic Program	(M = .575) $\tilde{M}$ = .575 N = 716 WN = 59,960	(M = .584) $\tilde{M}$ = .578 N = 236 WN = 24,076	(M = .636) $\tilde{M}$ = .596 N = 302 WN = 481,068	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants. The variable representing credits earned relative to normal progression is defined in Appendix F; these data were trimmed by excluding cases with suspect values (i.e., values of the variable greater than 1.5). Four entries are provided in each cell: (1) raw weighted means, M (in parentheses), (2) adjusted weighted means, $\tilde{M}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Excluding graduate study.

b/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

c/ Two a priori contrasts were formed from the adjusted means within each row: (1) column one versus column two, and (2) column one versus column three.

d/ Cell sizes in this row are sufficiently small to ensure that within stratum assumptions of the analytic model are violated; therefore, results should be viewed with extreme caution.

Table H.69

Postsecondary Graduation Rates as a Function of Upward Bound Participation Pattern And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a/]	Upward Bound Participation Pattern[b/]			Significance Level Of Contrasts[c/]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d/]	(P = .286) P̃ = .286 N = 90 WN = 6,394	(P = .528) P̃ = .498 N = 78 WN = 6,965	(P = .487) P̃ = .428 N = 134 WN = 128,314	(1) NS (2) NS
Two Year Academic Program	(P = .085) P̃ = .085 N = 216 WN = 14,752	(P = .104) P̃ = .091 N = 135 WN = 10,547	(P = .102) P̃ = .051 N = 223 WN = 207,718	(1) NS (2) NS
Four Year or Greater Academic Program	(P = .182) P̃ = .182 N = 846 WN = 55,761	(P = .186) P̃ = .186 N = 286 WN = 21,056	(P = .098) P̃ = .136 N = 372 WN = 426,089	(1) NS (2) NS

NOTE: These analyses were conducted for PSE entrants only. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, P̃ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Excluding graduate study.

b/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical-pattern category includes all other patterns of UB participation.

c/ Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

d/ Small cell sizes in this row led to insufficient data for within stratum estimates in some strata; therefore, results should be viewed with caution.

Table H.70

Proportion of Four-Year College Graduates Entering Graduate Study
As a Function of Upward Bound Participation Pattern

Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
Typical	Atypical	Nonparticipant	
(P = .089) $\tilde{P}$ = .089 N = 148 WN = 8,079	(P = .159) $\tilde{P}$ = .150 N = 54 WN = 3,311	(P = .090) $\tilde{P}$ = .117 N = 27 WN = 31,481	(1) NS (2) NS

NOTE: These analyses were restricted to the twelfth-grade cohort and to those individuals who had entered PSE and graduated from a four-year institution. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN. Cell sizes are sufficiently small to ensure that within stratum assumptions of the analytic model are violated; therefore, results should be viewed with extreme caution.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions: (1) column one versus column two, and (2) column one versus column three.

Table H.71

Postsecondary Educational Progress Relative to Goals
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering of Institution Last Attended[a]	Upward Bound Participation Pattern[b]			Significance Level Of Contrasts[c]
	Typical	Atypical	Nonparticipant	
Two Year or Less Vocational, Trade, or Business School[d]	($\bar{M}$ = 41.0%) $\tilde{M}$ = 41.0% N = 85 WN = 7,610	($\bar{M}$ = 34.6%) $\tilde{M}$ = 34.4% N = 72 WN = 9,437	($\bar{M}$ = 40.2%) $\tilde{M}$ = 32.5% N = 119 WN = 168,348	(1) NS (2) NS
Two Year Academic Program	($\bar{M}$ = 55.8%) $\tilde{M}$ = 55.8% N = 197 WN = 16,664	($\bar{M}$ = 54.3%) $\tilde{M}$ = 53.8% N = 130 WN = 14,375	($\bar{M}$ = 48.1%) $\tilde{M}$ = 50.0% N = 203 WN = 336,454	(1) NS (2) NS
Four Year or Greater Academic Program	($\bar{M}$ = 56.4%) $\tilde{M}$ = 56.4% N = 765 WN = 60,490	($\bar{M}$ = 56.3%) $\tilde{M}$ = 55.0% N = 254 WN = 26,427	($\bar{M}$ = 56.2%) $\tilde{M}$ = 58.1% N = 342 WN = 562,354	(1) NS (2) NS

NOTE: These analyses were restricted to PSE entrants; the variable reported is actual postsecondary progress expressed as a percentage of progress necessary to achieve reported educational goal (see Appendix E). Four entries are provided in each cell: (1) raw weighted means, $\bar{M}$ (in parentheses), (2) adjusted weighted means, $\tilde{M}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Excluding graduate study.

b/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

c/ Two a priori contrasts were formed from the adjusted means within each row: (1) column one versus column two, and (2) column one versus column three.

d/ Small cell sizes in this row led to insufficient data for within stratum estimates in some strata; therefore, results should be viewed with caution.

Table H.72

Rates of Attending Four-Year Institutions Within Twelfth-Grade Cohort
As a Function of Upward Bound Participation Pattern

Upward Bound Participation Pattern[a]			Significance Level Of Contrasts[b]
Typical	Atypical	Nonparticipant	
(P = .743) $\tilde{P}$ = .743 N = 610 WN = 39,113	(P = .654) $\tilde{P}$ = .658 N = 235 WN = 16,076	(P = .533) $\tilde{P}$ = .426 N = 285 WN = 229,326	(1) $p < .05$ (2) $p < .01$

NOTE: These analyses were restricted to PSE entrants in the twelfth-grade cohort. Four entries are provided in each cell: (1) raw weighted proportions, P (in parentheses), (2) adjusted weighted proportions, $\tilde{P}$ (adjusted within row for sex, race, poverty status, academic risk status, and over age in grade characteristics of the typical UB participation group), (3) cell size, N, and (4) total cell weight, WN.

a/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

b/ Two *a priori* contrasts were formed from the adjusted proportions within each row: (1) column one versus column two, and (2) column one versus column three.

Postsecondary Educational Progress of Twelfth-Grade Cohort
As a Function of Upward Bound Participation Pattern
And Level of Offering of Institution Last Attended

Level of Offering Of School Last Attended[a]	Postsecondary Educational Progress[b]	Upward Bound Participation Pattern[c]: Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[d]
Two Years Or Less[e]	Less Than One Year's Credit	(P = .644) $\tilde{P}$ = .644	(P = .711) $\tilde{P}$ = .812	(P = .515) $\tilde{P}$ = .738	(1) NS (2) NS
	One Year's Credit Or More	(P = .214) $\tilde{P}$ = .214	(P = .176) $\tilde{P}$ = .135	(P = .266) $\tilde{P}$ = .166	(1) NS (2) NS
	Graduated	(P = .142) $\tilde{P}$ = .142	(P = .113) $\tilde{P}$ = .053	(P = .219) $\tilde{P}$ = .096	(1) NS (2) NS
	Number of Cases	91	39	70	
	Total Weight	8,147	3,805	73,841	
Four Years	Less Than One Year's Credit	(P = .239) $\tilde{P}$ = .239	(P = .284) $\tilde{P}$ = .272	(P = .146) $\tilde{P}$ = .101	(1) NS (2) $p < .05$
	One to Two Years' Credit	(P = .175) $\tilde{P}$ = .175	(P = .162) $\tilde{P}$ = .164	(P = .139) $\tilde{P}$ = .102	(1) NS (2) NS
	Two to Three Years' Credit	(P = .112) $\tilde{P}$ = .112	(P = .084) $\tilde{P}$ = .121	(P = .236) $\tilde{P}$ = .308	(1) NS (2) NS
	Three Years' Credit or More	(P = .173) $\tilde{P}$ = .173	(P = .134) $\tilde{P}$ = .112	(P = .220) $\tilde{P}$ = .317	(1) NS (2) $p < .05$
	Graduated	(P = .301) $\tilde{P}$ = .301	(P = .335) $\tilde{P}$ = .332	(P = .259) $\tilde{P}$ = .172	(1) NS (2) NS
	Number of Cases	353	120	109	
	Total Weight	29,695	11,897	145,134	

NOTE: These analyses were restricted to PSE entrants in the twelfth-grade cohort. Two entries are provided in each cell: (1) raw weighted column proportions within major level-of-offering breakdown, P (in parentheses), and (2) adjusted weighte column proportions $\tilde{P}$ (adjusted within row for grade, sex, race, poverty status, academic risk status, and over age in grad characteristics of the typical UB participation group).

a/ Excluding any graduate training.

b/ One term's credit was determined as the equivalent of 45 quarter hours or 30 semester hours.

c/ Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

d/ Two *a priori* contrasts were formed from the adjusted column proportions or means within each row: (1) column one versus column two, and (2) column one versus column three.

e/ Cell sizes within this level-of-offering breakdown are sufficiently small to ensure that within stratum assumptions of the analytic model are violated; therefore, results should be viewed with extreme caution.

Appendix I

Procedures and Results of Relational Analyses

Appendix I

Procedures and Results of Relational Analyses

I. INTRODUCTION

This appendix addresses the relational analyses leading to the findings reported in Chapter 4 of the report. Specifically emphasized are: (a) the general analytic approach and rationale, (b) constraints and limitations on the analyses, (c) specific approaches to model building and testing, (d) analytic procedures employed, and (e) supporting tabular results. Topics that are closely related to these analyses are detailed elsewhere: weighting adjustments in Appendix G and analysis variable construction in Appendix E.

Although the comparative/descriptive analyses, described in Appendix H and reported in Chapter 3, address important questions of national impact of the Upward Bound (UB) program, there are other obviously important questions concerning the program and its possible impact at a less global level. Specifically, there is a need for a comprehensive approach to analysis in which student outcomes are examined in relation to the student characteristics (e.g., age, sex, ethnicity, poverty status, academic risk status), and grade level or contextual influences (e.g., use of PSE services, financial aid, institutional setting), as well as to the pattern of UB participation. The comparative/descriptive analyses suggested that differences in PSE outcomes among UB participants and nonparticipants may be explained through intervening factors of the PSE experience; however, the statistical procedures used did not permit simultaneous adjustment for a large number of intervening factors.

The relational analyses, which were conducted for a subset of outcome variables, represent an attempt to provide a comprehensive examination, through the formulation of a model of PSE outcomes and the examination of the relationships among the elements of that model. This modeling approach, embodied in recently developed techniques collectively known as analysis of covariance structures, provided a better means to exploit the multivariate nature of the data and to examine the complex structure of relationships among key variables. In general, this approach implies questions related to a fuller understanding of the PSE environments within which UB participants and nonparticipants are placed.

These analyses start with a model of how various data elements relate to each other over time, i.e., a model of directional relationships. Within the adopted model, directional prediction paths are specified for all elements, and the overall relationship ("effect") between two variables is decomposed into (a) a part that operates through intermediate variables, indirect effects, and (b) a residual part that cannot be attributable to intermediate factors, the direct effect. Thus, some portion of total relationship is due to relationships through other variables and the remaining portion represents the direct relationship that is not otherwise explained.

The technique is particularly suited to the longitudinal nature of this study, but is less well suited to the underlying sample methodology. While the developed model is convenient in explaining the data and intuitively appealing, there are probably other models that could fit the data as well or better. Moreover, the technique used required considerably more assumptions than the hypothesis-testing approach, and results must be considered exploratory and should be interpreted with considerable caution.

The remaining sections of this appendix provide greater detail concerning the procedures used and results of this set of exploratory analyses. Section II presents general analytic and modeling considerations. Section III details the specific analytic technique used, and Section V provides details of additional methodological procedures that were implemented in the specific application of the analytic technique to the current study. Section V provides a discussion of the several constraints and limitations of these analyses. Section VI presents major analytic results in tabular form.

II. GENERAL CONSIDERATIONS

A. Analytic Considerations

Perhaps the most widely known technique of structural equation approaches is path analysis,[1] which has been described as "a method for studying the direct and indirect effects of variables taken as causes on variables taken as effects."[2] The procedure used in these analyses is quite similar to path analyses, but but actually represents a more general procedure developed for analysis of covariance structures.[3] The use of the terms "causes" and "effects" is somewhat misleading since the approaches do not examine cause in a metaphysical sense but investigate potential causation in a predictive or explanatory sense.

Application of the approach is in testing a model of directional relationships, formulated on the basis of prior knowledge and/or theoretical considerations. Thus, structural equation models (and solutions) involving the same constructs can vary with the tenets of the particular theoretical framework adopted. Moreover, the inclusion or exclusion of particular variables from the model can change solutions dramatically. Analytic solutions indicate whether the model examined is plausible (not necessarily the most plausible) and, the nature of the relationships within the assumed model. Thus, the techniques are directed toward examining data within a particular theory of behavior rather than in generating such theory.

Structural equation approaches investigate both the direct and indirect effects of predictor variables on criterion variables, in contrast to standard regression analysis, which assesses only the simultaneous direct unique relationships of variables to an outcome. The different types of effects are determined through a decomposition of the overall relationship of one variable to another into (a) the indirect effect, the part that operates only through other intermediate varables in the model, and (b) the direct effect, the residual portion of the relationship that is unique (i.e., does not operate through intermediate

---

1 See Duncan, O. D., Introduction to Structural Equation Models. New York: Academic Press, 1975.

2 Kerlinger, F. N. and E. J. Pedhazur, Multiple Regression in Behavioral Research. New York: Holt, Rinehart & Winston, 1973.

3 Jöreskog, K. G. A General Method for Analysis of Covariance Structures, Biometrika, 57(2), 1970, p. 239-251.

variables). Thus, within the adopted model, it is assumed that some of the overall relationship of one variable to another is due to relationships (links) that exist with other variables that intervene and a residual portion of the relationship that cannot be explained through the intervening variables.

Structural models can be expressed by means of directional paths interrelating the sets of variables considered, and solutions typically yield coefficients associated with these paths, which may be interpreted in a similar manner to regression coefficients. The coefficient associated with the path that leads directly from a variable of interest to the criterion variable considered represents the direct effect portion of the relationships, while the sum of the products of the coefficients associated with paths that lead through intervening variables may be taken as the indirect effect.

B. Modeling Considerations

The model examined through structural equation approaches should be based in theory, since the posited model is thereafter assumed to be the only representation of the structural interrelationships that is considered. This implies, of course, that all results are conditional on the model that has been assumed. Several models were considered, in light of their theoretical soundness, the availability and scale properties of key variables, and potential technical problems. The nature of the model finally adopted is depicted in Figure I.1.

The adopted model posits specific directional relationships (paths) among the elements; the direction of a relationships is indicated by an arrow. Two sets of variables are taken as starting points for prediction: UB participation pattern and background characteristics. These starting-point variables are called exogenous variables. Relationships are assumed to exist between the two classes of exogenous variables and among the exogenous variables within each set. There is, however, no clearcut directionality among the exogenous variables on the basis of theory; thus, no direction is assumed for the relationships.

Both of the exogenous variable sets are considered to predict subsequent variable sets within the model, which are called endogenous variables. As indicated in Figure I.1, additional prediction is allowed among the ordered endogenous variable sets. The directionality of prediction indicated in the figure represents a major modeling assumption.

The first prediction link is to the set of variables representing PSE context factors. The next prediction links in the model are to the set of variables representing use of support services. This implies that the institutional context predicts service usage. Thus, prediction of service use variables is assumed from UB participation pattern, background characteristics, and from PSE context factors. Indirect effects on service use may exist through PSE context factors (e.g., the relationship between UB participation pattern and use of support services may be partially explained through the relational link of UB participation to PSE context factors and the subsequent link of context factors to use of services). The class of financial aid variables is modeled as related to all previously considered factors, and PSE success outcomes are assumed to be related to all other elements in the model.

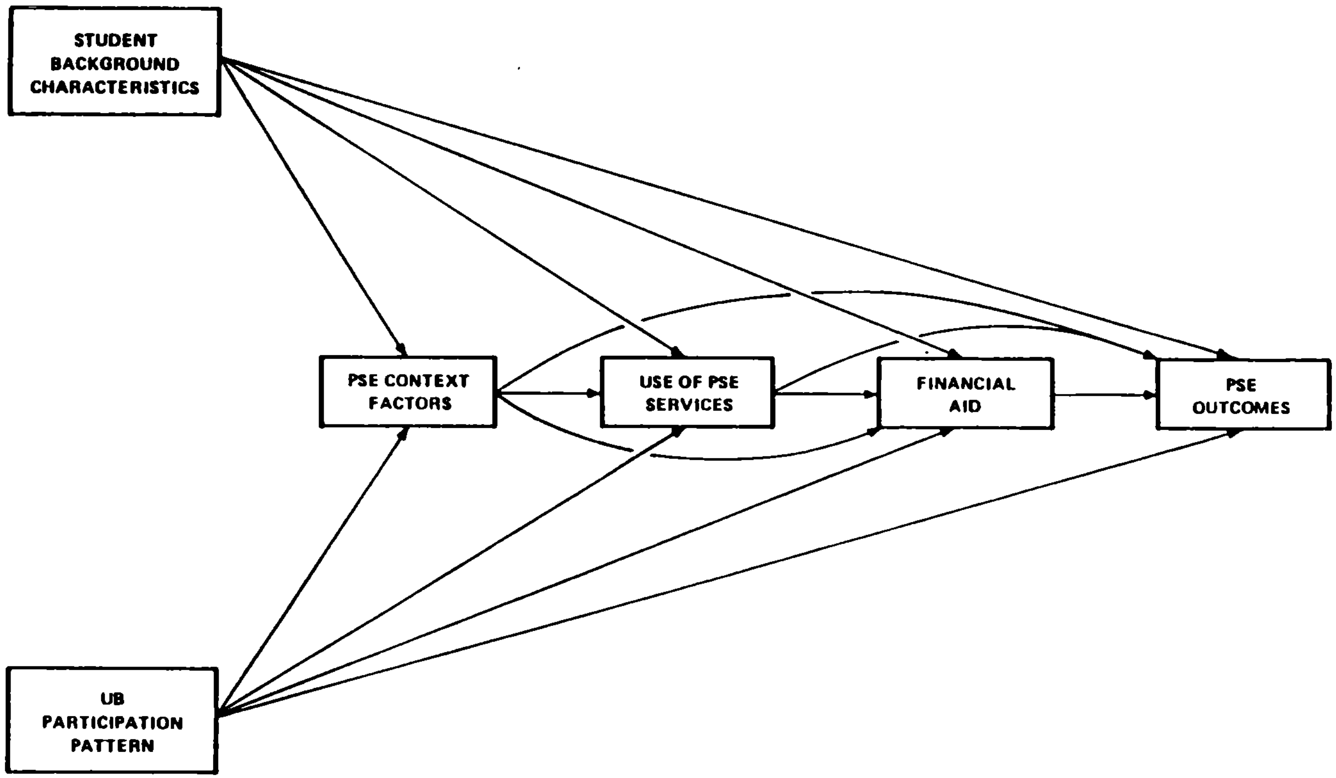


Figure I.1 General form of the adopted theoretical model of PSE outcomes.

Even within the simplified conceptual model presented in Figure 16, there are 11 potential direct effects and 22 potential indirect effects; however, most of the elements in the simplified model represent more than one variable. Thus, the actual model considered represented a much more complex structure of relationships. For the 24 variables that were considered (see Section IV), the number of potential direct and indirect effects was greatly increased. Since each variable set indicated in Figure I.1 represents more than one variable, it was also necessary to model the relationships among variables within each set. Directionality within a set typically was neither theoretically nor logically indicated, and such relationships were modeled as nondirectional, allowing the necessary existence of the relationships but introducing links through which indirect effects could operate that were quite difficult to interpret. Regardless of the complexity, however, the basic assumptions of the model (e.g., use of PSE services precedes and thus directionally predicts receipt of financial aid and PSE success measures) are defined by Figure I.1.[4]

III. THE ANALYTIC MODEL

The analysis of linear structural relationships by the method of maximum likelihood is an approach to the analysis of the interrelationships among variables which has been under development by Karl Jöreskog and others over the past decade.[5] The purpose of the approach is to permit the estimation of parameters in models relating a set of exogenous variables to a set of endogenous variables within which additional relationships may be of interest. With the proper restrictions, this approach will yield the same results as an estimation of path coefficients through the use of regression analysis, an approach widely used by sociologists.[6] The approach can also be used appropriately where variable measurement is fallible, a situation in which regression analysis yields biased estimates of parameters. Given multiple measures of each fallibly measured variable, the method will even permit the simultaneous estimation of factor loadings, measurement error, and the interrelationships among exogenous and endogenous latent variables. In this sense, the method permits investigators to explore more characteristics of their constructs than other avaiable methods.

Since multiple measures of the variables of interest were not readily available, the present application of this method does not make use of the factor analytic feature; however, in the second set of relational analyses considered, estimates of error variance associated with each measure were drawn from other components of the study to permit an investigation of interrelationships among the latent variables represented by the measures.

---

[4] To the extent that these basic assumptions are invalid, results may represent little more than the end product of a computational exercise.

[5] For additional references to work in this area, see Jöreskog, K. G. and Sörbom, D. <u>LISREL: Analysis of Linear Structural Relationships by the Method of Maximum Likelihood, User's Guide (Version IV, Release 2)</u>. Chicago, Ill.: National Educational Resources, Inc., November 1978.

[6] Duncan, O. D. (<u>op</u>. <u>cit</u>.)

A. The Basic Model Without Measurement Error

Since the sufficient statistic for the use of this method is the variance-covariance matrix among the complete set of exogenous and endogenous variables, the simplest statement of the model is in terms of this matrix. Given both exogenous and endogenous variables, it is convenient to partition the overall variance-covariance matrix among the variables as follows:

$$\Sigma = \left[\begin{array}{c|c} \Sigma_{yy} & \Sigma_{yx} \\ \hline \Sigma_{xy} & \Sigma_{xx} \end{array}\right]$$

where

Σ_{yy} is the variance-covariance matrix among the endogenous variables,

Σ_{xx} is the variance-covariance matrix among the exogenous variables,

and $\Sigma_{yx} = \Sigma_{xy}'$ is the variance-covariance matrix between the exogenous and endogeous variables.

For the model involving no measurement error, each of these components can be expressed in terms of simpler parameter matrices, estimates of which are of practical interest. Using the notational scheme employed by Jöreskog the parameterized expression for the overall variance-covariance matrix can be expressed as:

$$\Sigma = \left[\begin{array}{c|c} \beta^{-1} (\Gamma \Phi \Gamma' + \Psi)\beta^{-1'} & \beta^{-1} \Gamma \Phi \\ \hline \Phi \Gamma' \beta^{-1'} & \Phi \end{array}\right],$$

where

Φ is the variance-covariance matrix among the exogenous variables,

Γ is a matrix of coefficients relating the exogenous to the endogeous variables,

β is a matrix of coefficients relating the endogenous variables to one another, and

Ψ is a matrix representing the residual variance-covariance matrix among the endogenous variables unaccounted for by the exogenous variables.

The model may be more easily understood in terms of a simple example. Assume two exogenous variables (x_1 and x_2) and two endogenous variables (y_1 and y_2). A simple path diagram representing interrelationships among these variables could be expressed as shown in Figure I.2. In this particular model, the exogenous variables, x_1 and x_2, are seen as mutually interrelated predictors of the first endogenous variable, y_1. The latter variable, in

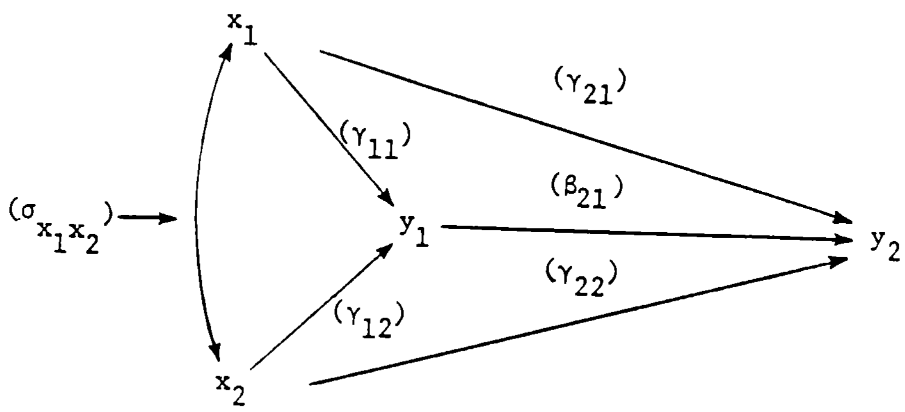


Figure I.2. Hypothetical Model 1.

concert with the exogenous variables, is seen as predicting the second endogenous variable, y_2. The coefficient parameter matrices in this case would be:

$$\Gamma = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}, \qquad \beta = \begin{bmatrix} 1 & 0 \\ -\beta_{21} & 1 \end{bmatrix}.$$

The remaining matrices would be defined as:

$$\Phi = \begin{bmatrix} \sigma_{x_1}^{2} & \sigma_{x_1x_2} \\ \sigma_{x_2x_1} & \sigma_{x_2}^{2} \end{bmatrix}, \qquad \Psi = \begin{bmatrix} r_{y_1}^{2} & r_{y_1y_2} \\ r_{y_2y_1} & r_{y_2}^{2} \end{bmatrix},$$

where $\sigma_{x_1}^2$ and $\sigma_{x_2}^2$ are the variances for x_1 and x_2 respectively,

$\sigma_{x_1x_2} = \sigma_{x_2x_1}$ is the covariance among x_1 and x_2,

$r_{y_1}^2$ and $r_{y_2}^2$ are the residual variances for y_1 and y_2, and, finally,

$r_{y_1y_2} = r_{y_2y_1}$ is the residual covariance among y_1 and y_2.

A very powerful aspect of the approach is that parameters can be constrained equal to zero, thereby permitting the estimation of simpler models such as the one shown in Figure I.3. Here the x's no longer directly influence y_2, but operate indirectly through y_1. Using a likelihood ratio approach, the "fit" of Model 2 can be compared statistically with that of Model 1. The absence of a significant difference would imply that Model 2 is as statistically valid as Model 1. Given the assumed a priori directionality among the x's and the y's, this result could be taken as supporting an interpretation that x_1 and x_2 predict y_1, which, in turn, is the sole direct predictor of y_2 (i.e., all prediction of y_2 by x_1 or x_2 is indirect through y_1). Of course, other simplifications of Model 1 are possible.

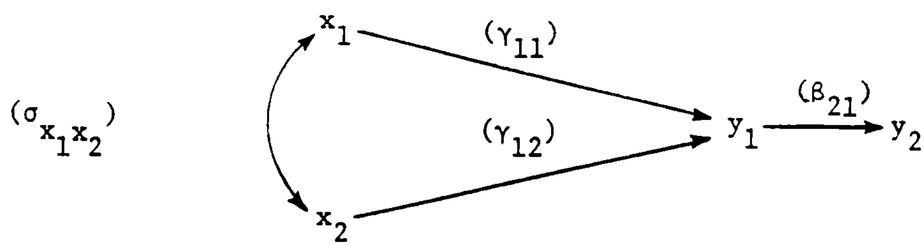


Figure I.3. Hypothetical Model 2.

B. <u>The Model with Measurement Error</u>

Given known measurement error for the variables employed, the resulting model may be expressed as:

$$\Sigma = \left[\begin{array}{c|c} \beta^{-1}(\Gamma\Phi\Gamma' + \Psi)\beta^{-t} + \Theta_\varepsilon & \beta^{-1}\Gamma\Phi \\ \hline \Phi\Gamma'\beta^{-t} & \Phi + \Theta_\omega \end{array}\right],$$

where Θ_ε and Θ_ω are diagonal matrices containing the estimated error variance associated with the endogenous and exogenous variables, respectively. This is a more general model, which reduces to the previous case when $\Theta_\varepsilon = 0$ and $\Theta_\omega = 0$ (i.e., no measurement error). Under these circumstances, however, the meaning of the other parameter matrices is somewhat changed, as follows:

Φ is the variance-covariance matrix of the "true" values of the exogenous variables, Γ is a matrix of coefficients relating the "true" values of the exogenous variables to the "true" values of the endogenous variables,

β is a matrix of coefficients relating the "true" values of the endogenous variables to one another, and

Ψ is the variance-covariance matrix of the residual "true" values of the endogenous variables.

The estimated coefficients in Γ and β, therefore, refer to the interrelationships among the latent variables and not the fallibly measured observed variables.

In terms of Models 1 and 2 discussed above, the inclusion of the measurement error components generates the modified path models shown in Figures I.4 and I.5. In the modified models, ω_1 and ω_2 refer to the measurement error in x_1 and x_2 respectively, ε_1 and ε_2 refer to the measurement error in y_1 and y_2 respectively, and the other coefficients (not shown) are represented as previously indicated.

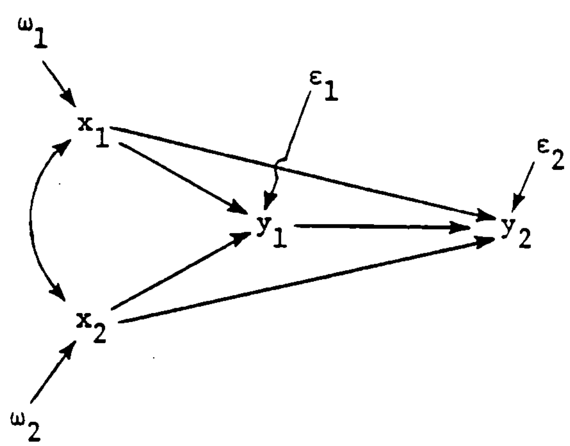


Figure I.4. Modified hypothetical Model 1.

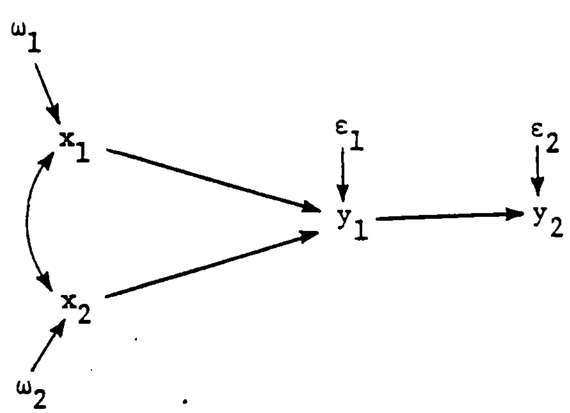


Figure I.5. Modified hypothetical Model 2.

C. Parameter Estimation, Model Testing, and Standard Errors

Given a variance-covariance matrix and a model containing parameters to be estimated, the approach generates estimates for which the likelihood of the observed variance-covariance matrix is maximized under the assumption that the underlying data is distributed multivariate-normal. Standard errors for the estimated parameters are produced by evaluating the negative inverse of the information matrix at the estimates. The square roots of the diagonal values of the resulting matrix provide asymptotic standard errors for the estimates.

An overall statistical test of fit of a given model is obtained by taking the ratio of the likelihood under the estimated model to the likelihood under a completely unrestricted model where Σ is estimated by the observed variance-covariance matrix. For large samples and under the null nypothesis that the estimated model fits the data, this statistic is distributed as chi-square. A parallel procedure can be employed to statistically evaluate the fit of one estimated model as compared to that of another.

IV. SPECIFIC ANALYTIC PROCEDURES

The specific software used in solving the posited structural model was the LISREL program.[7] This software is particularly applicable to the present problem. First, the general procedure allows for the posited nondirectional relationships within the separate classes of intervening variables, which cannot be accomodated by many approaches to the problem. Second, the procedure allows investigation of differential variable reliability (which is suggested for the data used), another feature that cannot be accomodated by simpler regression approaches.

To implement both the general procedure and the specific software, detailed decisions were necessary regarding the exact nature of the modeling, weighting, and variables. Moreover, within the developed model, specific approaches to the solutions reported in Chapter 4 of the report were not always completely straightforward. A description of these study-specific procedures is provided in the subsections that follow.

A. Additional Modeling

Most of the modeling was relatively uncomplicated. Paths were allowed from all variables in the variable sets defined in Figure I.1 to all variables in subsequent sets, which defined the matrices Γ and β (described in Section III). While nondirectional relationships among the exogenous variables were easily incorporated in the matrix Φ, the nondirectional relationships of endogenous variables within the same variable set were more problematic. Most structural equation approaches require specification of directionality for all endogenous variables (which would allow the relationships to be represented in β); however, in most cases no directionality within variable set could be specified, for the model considered here.

These nondirectional relationships among endogenous varibles were allowed to be represented in the matrix Ψ, which contains residual variances and covariances among the endogenous variables. Under typical modeling the covariance of errors of prediction would be constrained to zero (i.e., Ψ is usually a diagonal matrix), but the representation used here allowed the nondirectional relationships to be represented through correlated errors of prediction. This

---

7 Jöreskog, K. G. and Sörbom, D., November 1978 (op. cit.)

representation is considered theoretically sound, but introduced some problems in interpretation (see below, Subsection IV.E).

For the variable set involving financial aid elements, directionality of the relationships among the variables was assumed on the basis of a model recommended by USOE,[8] which was of sufficient intuitive appeal to warrant consideration. In this modeling, the amount of grant received (as rescaled to account for differential institutional costs) was taken as the primary form of financial aid, on which receipt or use of other financial aid elements depended. The modeling assumed that, given receipt of some amount of grant/scholarship funds, the next source of aid sought was loans and, following that, earnings or savings were used to round out the financial aid package. These directional relationships were incorporated into the matrix β.

For modeling involving differential reliabilities, another problem occurred in representing measurement error in the matrices Θ_ε and Θ_ω, and associated problems in modeling the diagonal elements of Φ. Since multiple measures of the variables were not available, the program could not provide these estimates directly; rather, the values had to be obtained from other study components. Where reliability (r) estimates were available, error variances (s_e^2) were computed from observed variances (s_o^2) through the well-known result of classical test theory:

$$s_e^2 = (1 - r)\, s_o^2.$$

Where reliability estimates were not available, the values of error variance used were determined jointly by RTI and USOE, based on available information and subjective estimates. Given estimates of error variance, the true score variance (s_t^2) estimates for Φ were obtained from another result of test theory:

$$s_t^2 = s_o^2 - s_e^2.$$

B. Weighting

Weighting of the data for analysis posed particularly difficult problems. A large number of technical problems were introduced through attempts to use sampling weights, as adjusted. Former UB participants and nonparticipants were represented by markedly different sampling weight sums (reflecting the markedly different sizes of the populations from which they were drawn). Moreover, there were known differences between the former participant and nonparticipant groups in the distributions of background factors (see Table H.1 in Appendix H). While the background factors were incorporated into the analytic model, the differential distributions coupled with the weight differentials could have resulted in relationships among exogenous and endogenous variables that were determined almost exclusively by the relationships existing in the nonparticipant group.

---

8 The directionality assumed was similar to that investigated by Carroll. See Carroll, C. D. Packaging of Financial Aid for First-Time, Full-Time Dependent Freshmen in 1978. Technical Paper 79-13. Washington, D.C.: Office of Evaluation and Dissemination, U.S. Office of Education, August 3, 1979.

To overcome this problem, RTI and USOE agreed upon a weighting procedure that was unrelated to sampling rates. The approach used was based on a desire to create a quasiexperimental situation in whch outcomes could be modeled for two more-or-less equivalently sized groups with analogous distributions of background factors. Since individuals for whom the entire set of modeling variables was not available (i.e., item nonresponse problems) could not be used in the analyses, the weighting was performed only for the subset of individuals with complete data. The procedure was a technique whereby two artificial groups were constructed such that the nonparticipant group would be similar to the UB participant group on relevant student background variables (sex, race, poverty status, academic risk status, and initial grade cohort distribution) and in total group size. In essence, the procedure forced a greater similarity between the two groups in terms of the background characteristics that are theoretically related to outcomes but that were not completely controlled by the study design.

This adjustment procedure (a form of partial balancing) was used exclusively for the relational analyses, and is detailed in Appendix G. It should be noted that the balancing procedure used in this study (like that used in previous studies) only attempted adjustment for two groups--former participants and nonparticipants. Adjustment within the former participant group on the basis of participation pattern was not attempted. Resulting distributions of background characteristics for former participants and nonparticipants using these weights are given in Table I.1. The procedure was not completely successful; it overadjusted the nonparticipant group for poverty status, race, and size. On other factors, however, the procedure produced comparable marginal frequencies for the variables considered.

C. Variable Construction

The analytic approach is predicated on the use of continuous variables with at least interval scale properties; however, many of the raw variables were categorical and others represented only an ordinal scaling. The nature of the variables constructed for analysis was determined jointly by RTI and USOE and, while not perfect, represents an attempt to produce the most meaningful analytic variables from the available data.

Each of the classes of variables shown in Figure 16 was represented in the full model by one or more variables. Those variables that were categorical in nature were coded as (0,1) variables to be more amenable to relational analysis, yielding a base group (coded 0) and a predictor group (coded 1). In this case, resulting coefficients reflect the effect associated with the predictor group relative to the base group (e.g., the effect associated with being female rather than male). Some categorical variables represent three classes and must be dealt with differently. The approach used in this case was dummy coding, which creates two related dichotomies, in which case the group coded 0 in both dichotomies was the base group. Although some assumptions are violated in using such variables, the procedure is commonly used and generally acceptable.

All participation pattern variables and background variables were represented by a recoding of categorical raw variables. The resulting analytic variables are defined in Table I.2. Base groups for these variables are obvious from the coding. The five PSE contextual factors were coded or scaled

Table I.1

Demographic Characteristics of the Relational Analysis Subgroup After Partial Balancing
As a Function of Upward Bound Participation Pattern

Variable[a]	Classification	Upward Bound Participation Pattern[b]: Typical	Atypical	Nonparticipant	Significance Level Of Contrasts[c]
Sex[d]	Male	.445	.410	.367	(1) NS
	Female	.555	.590	.623	(2) NS
Age Within Grade[d]	Not Over Age In Grade	.931	.893	.898	(1) NS (2) NS
	Over Age In Grade	.069	.107	.102	
Poverty Status[d]	Poverty Level	.765	.628	.866	(1) p < .01
	Not Poverty Level	.235	.372	.134	(2) p < .01
Academic Risk Status[d]	Risk	.367	.341	.465	(1) NS
	Not Risk	.633	.659	.535	(2) NS
Race	Black	.721	.567	.770	(1) p < .01 (2) NS
	White	.117	.234	.071	(1) p < .01 (2) p < .05
	Other	.162	.199	.159	(1) NS (2) NS
Grade Cohort[e]	10th Grade	.104	.134	.109	(1) NS (2) NS
	11th Grade	.401	.402	.405	(1) NS (2) NS
	12th Grade	.495	.464	.486	(1) NS (2) NS
Number of Cases		656	261	376	
Weight Total		8,528	3,393	16,542	

NOTE: Cell entries are weighted column proportions within major variable categories, based on equal weights for former participants and with nonparticipant weights partially balanced on marginal distributional differences in the unweighted participant and nonparticipant groups, and further adjusted to obtain more equivalent group sizes. All adjusted weights were multiplied by a constant of 13 to ensure that no weight would be less than 1.

[a] These variables determined during base year study.

[b] Typical UB participation pattern is defined as those participating in grades 10, 11, and 12 or in grades 11 and 12; the atypical pattern category includes all other patterns of UB participation.

[c] Unless otherwise indicated, two *a priori* contrasts were formed from the column proportions within each row: (1) column one versus column two, and (2) column one versus column three.

[d] Only one set of contrasts is reported for this variable, since the contrasts for each of the two classifications of the variable are completely redundant.

[e] Represents grade level during academic year 1973-74.

Table I.2

DUMMY CODING OF THE EXOGENOUS VARIABLES USED IN MODELING

Original Variable	Values of Original Variable	First Dummy Variable	Values of First Dummy Variable	Second Dummy Variable	Values of Second Dummy Variab
UB Participation Pattern	0 = Nonparticipant 1 = Atypical Participant 2 = Typical Participant	UBTYP	0 0 1	UBATYP	0 1 0
Grade Cohort	1 = 10th Grade 2 = 11th Grade 3 = 12th Grade	GR10	1 0 0	GR12	0 0 1
Sex	1 = Male 2 = Female	SEX	0 1		
Academic Risk	1 = Academic Risk 2 = Not Risk	RISK	1 0		
Race	1 = Black 2 = White 3 = Other	RABL	1 0 0	RAOTH	0 0 1
Poverty Status	1 = Poverty Level 2 = Not Poverty Level	POVS	1 0		

NOTE: The dummy variables correspond to the values of the original variables given on the same lines.

as indicated in Table I.3. The host status variables represent a recoding in which non-host institutions were the base group; the remaining three variables were considered to represent scaled values. It should be recognized that, while the constructs underlying selectivity and size may be represented on continuous interval scales, the actual variables used are transformations of the scales that preserve only the order. The level-of-offering variable has, at best, only ordinal scale properties. Three support service use variables were considered, representing extent of using services within the three general support areas of tutoring, counseling, and remedial services. Each of these variables was scaled as indicated in Table I.4, but these scales also have only ordinal properties.

Four meaures of financial aid were considered: receipt of grants or scholarships (GRNTR), (2) receipt of loans (LOANR), (3) use of earnings or savings (SWORKR), and (4) scaled amount of grants/scholarhips received (GRTCAT). The first three variables resulted from a recoding with nonreceipt representing the base group. The final measure represents a five-valued scaling of total amount of grant/scholarship funds received, as adjusted by the cost of the institution last attended. Preliminary trimming of the grant/scholarhip amount variable is shown in Tables I.5 and the scaling of GRTCAT for values other than zero is defined in Table I.6. The final scale has only ordinal properties.

Three PSE success outcomes were considered: (1) a persistence measure, number of terms of attending PSE (NOTRM), (2) a performance measure, overall GPA at instituion last attended (TOTGPA), and (3) a progress measure, total termwise credits earned (RTCRED). These variables are considered to represent interval or ratio scales. The first variable is scaled as a whole number from 0 to 25; the second variable is continuous from 0.0 to 4.0. The last variable is a rescaling of total credits earned by dividing by 15 (the number of credits needed per term to progress normally in student class), which maintains the ratio scale property of the variable.

D. Solution Strategies

The solution of the model was not completely straightforward. Cost considerations necessitated implementing a stagewise solution using some preliminary modeling, rather than a simultaneous solution for all potential model parameters. Additionally, the nondirectional relationships among variables within certain endogenous variable sets posed a very real problem in the computation of indirect effects; thus, alternate strategies were implemented to estimate these effects. The particular strategies used are detailed below.

1. Preliminary Modeling

The full model proposed involved 9 exogenous variables and 14 endogenous variables. With the adopted specification of the model, a completely unconstrained solution would have involved estimating 231 free parameters for the case without measurement error and additional free parameters in the estimated measurement error case. The estimation procedure for the LISREL program is quite efficient for moderate-sized problems, using a sequential application of steepest descent and Fletcher-Powell algorithms. For larger problems, the solution time increases geometricaly with additional variables. Moreover, in extreme cases it is possible that convergence will not obtain

Table I.3

SCALING OF THE INSTITUTIONAL CHARACTERISTICS VARIABLES
USED IN MODELING

Variable	Categories	Scale Value
IOFF (Institutional level of offering)	Vocational/technical/business	-1
	Two-year college	0
	Four-year college or university	1
ISIZ (Undergraduate student enrollment)	Less than 500	-2
	500-2,499	-1
	2,500-7,499	0
	7,500-14,999	1
	15,000 or more	2
ISEL (Relative institutional selectivity)	Low	-1
	Intermediate	0
	High	1
UBHOST (UB Host status of institution 1974-1978)	Not host	0
	Host	1
SSHOST (SSDS host status of institution 1974-78)	Not host	0
	Host	1

Table I.4

SCALING OF THE SERVICE USE VARIABLES
USED IN MODELING: USE OF TUTORING SERVICES (TUSE),
USE OF COUNSELING SERVICES (CUSE),
AND USE OF REMEDIAL SERVICES (RUSE)

Condition	Scale Value
Not available or did not participate	0
Participated "infrequently"	1
Participated "sometimes"	2
Participated "frequently"	3

Table I.5

TRIMMING EXCESSIVELY HIGH OR LOW VALUES
OF AMOUNT OF GRANTS/SCHOLARHIPS RECEIVED

Institutional Costs	Old Value	New Value[a]
Any value	\$0	\$0
Any value	\$1-\$10	\$0
Any value	\$11-\$100	\$100 + old value
Any value	\$101-\$5,000	Old value
Less than \$500	\$5,001 or greater	R{OLD VALUE/10}
\$500-\$999	\$5,001-\$10,000	R{OLD VALUE/5}
\$500-\$999	\$10,001 or greater	R{OLD VALUE/10}
\$1,000-\$1,999	\$5,001-\$8,000	R{OLD VALUE/2}
\$1,000-\$1,999	\$8,001-\$12,000	R{OLD VALUE/3}
\$1,000-\$1,999	\$12,001 or greater	R{OLD VALUE/4}
\$2,000-\$3,999	\$5,001-\$7,500	R{OLD VALUE/1.5}
\$2,000-\$3,999	\$7,501-\$10,000	R{OLD VALUE/2}
\$2,000-\$3,999	\$10,001 or greater	R{OLD VALUE/3}
\$4,000 or more	\$5,001 or greater	\$4,000 + R{OLD VALUE/10}
Indeterminate	\$5,001 or greater	\$2,000 + R{OLD VALUE/10}

a/ R{x} = x rounded to nearest integer value.

Table I.6

SCALING OF AMOUNT OF GRANTS RECEIVED

Institutional Cost	Reported Amount of Grants/ Scholarships Received[a/]	Scale for GRTCAT
Less than $500	Less than $250	1
	$250-$499	2
	$500-$999	3
	$1,000 or more	4
$500-$999	Less than $500	1
	$500-$999	2
	$1,000-$1,999	3
	$2,000 or more	4
$1,000-$1,999	Less than $1,000	1
	$1,000-$1,999	2
	$2,000-$3,999	3
	$4,000 or more	4
$2,000-$3,999	Less than $2,000	1
	$2,000-$3,999	2
	$4,000-$7,999	3
	$8,000 or more	4
$4,000 or more	Less than $4,000	1
	$4,000-$7,999	2
	$8,000-$15,999	3
	$16,000 or more	4

a/ After trimming (see Table I.5).

within the maximum allowable time for computer usage. It was therefore considered infeasible to attempt a simultaneous solution for all potential free parameters. Rather, a stagewise approach to the model solution was implemented; the approach taken differed somewhat for the case without measurement error and the case with estimated measurement error.

In solving the model without measurement error, some relationships were constrained to zero on theoretical grounds (e.g., grade cohort relationships to institutional context variables) by agreement of RTI and USOE. Additionally, under these constraints, preliminary regression solutions were implemented to detect empirical nonsignificant relationships. Using a stepdown approach, each endogenous variable was examined in relation to all potential exogenous and endogenous variables that could predict it within the general framework of the adopted model. Predictor variables with nonsignificant regression coefficients were removed from the regression solution in a stepwise manner, using the magnitude of the F ratio as the criterion for sequential stepdown. Two or more dummy variables created from the same underlying categorical variable (i.e., grade cohort, UB participation pattern, race) were treated as a class variable. Thus, dummy variables in the same class were eliminated from the regression model only if the prediction of the entire class was nonsignificant. The preliminary regression approach resulted in parsimonious prediction models for each endogenous variable. Only the predictor variables remaining in each of the parsimonious models were considered in the LISREL solution; other variables that had been eliminated on theoretical grounds or through the regression solutions were constrained to values of zero in the matrices Γ and β.

Beyond this, the LISREL solution itself was obtained in a stagewise manner. Due to specific USOE interest, the grant/scholarship receipt variable was considered separately from the full model. Since the rescaled value of amount of grants/scholarship funds received was considered a more meaningful variable for the full model, it was used, along with receipt of loans and use of earnings or savings, in subsequent modeling of the financial aid variable complex and of postsecondary performance, progress, and persistence. Given the need for two model solutions, the simpler model involving grant/scholarship receipt was solved first; that solution yielded solutions for parameters of the relationships among exogenous variables, the set of institutional context variables, and the set of service use variables. These solutions were then used as fixed parameters in the subsequent solution of the full model. In both solutions, iterations were necessary due to the statistical nonsignificance of some coefficients in the first solution. The iterative approach used was analogous to that employed in the preliminary regression approach.

For the model with estimated measurement error, a priori constraints were not introduced, and the preliminary regression approach was not appropriate, since it could not accomodate the differential reliability estimates. In this solution a different stagewise approach was used. The approach was conceptually quite simple, involving a stepwise LISREL solution to each set of variables in the model. First a solution was obtained for the first set of endogenous variables, the institutional context factors. This solution started by allowing all possible relationships as free parameters, and an iterative parsimonious solution was obtained by constraining nonsignificant coefficients to zero in a stepwise manner. The parameters from the parsimonious solution for institutional context (including those that had been constrained to zero) were used as fixed parameters in modeling the next set of endogenous variables, service use. The process was continued until the last set of endogenous variables, PSE success indices, had been included.

2. Alternate Solutions for Indirect Effects

As indicated by the structural equations approach, direct effects of UB participation on the PSE success measures examined can be misleading without consideration of indirect effects through the three intervening variable sets (school type, service use, and financial aid). The numerical value of indirect effect can be obtained through multiplication of appropriate coefficients associated with the indirect paths as indicated in the tables in Section VI of this appendix. As indicated earlier, however, those coefficients do not reflect the additional indirect effects operating through the nondirectional paths among endogenous variables within a particular variable set (for example, the relationships among the service use variables). While such relationships have been allowed in the structural model solution, they are not reflected in the solution for path coefficients reported for the matrices Γ and β.

An alternate approach was therefore used to obtain these indirect effects for the PSE success variables. In general, the approach involved solving a structural model that did not involve the intervening variables, but that did include the background and temporal factors. The unique relationship between typical and atypical UB participation, as obtained through such an alternate solution, reflects the overall relationship between UB participation and the outcome variable of interest. The total indirect relationship was then obtained by subtracting the direct relationship, as indicated in Γ for the full modeling, from the total overall relationship. Moreover, assuming that the relative contribution of school type, service use, and financial aid to the total indirect effect was proportionally equivalent for the alternate solution and the direct solution (obtained through multiplication of the path coefficients), unique contributions of each of these three classes of variable on outcome measures of interest were estimated within the alternate approach.

The solution for the overall effect differed somewhat for the model without error and the model with estimated error. In the former case, solutions were obtained using a multiple regression approach, but in the latter case, in which it was necessary to account for differential reliability, a separate LISREL solution was obtained. The two solutions for overall effects are given in Section VI, Tables I.23 and I.24. Additional computations of indirect effects through the alternate approach are provided in Tables I.26 through I.31.

V. CONSTRAINTS AND LIMITATIONS OF ANALYSES

The relational analyses are subject to some of the same constraints and limitations as the comparative/descriptive analyses, as well as several others. The potential problems of sampling error and bias remain, as do other constraints involving data validity and interpretation (see Section II of Appendix H). Additional limitations associated specifically with the relational analyses can be classified into two areas: (a) limitations of the technique per se and (b) limitations specific to the use of the technique within the current study. Limitations of the first type are inherent in assumptions that underly the procedure, while limitations of the second type concern additional assumptions made in current use of the procedure.

A. Limitations of the Technique

The structural equations approach assumes a model that has been developed and, thus, is subject to errors in specification of that model. While tests are available to determine if the model provides a reasonable fit to the data, no tests are available to determine if the adopted model provides a "best fit." Moreover, all estimates produced from the analysis are completely conditional on the model that has been assumed (different models could produce different estimates). Consequently, results are considerably assumption-bound.

The major constraining assumptions are: (1) all relevant variables have been incorporated into the model, and (2) the model has been correctly specified. If these assumptions are met, then results are theoretically sound. There is no assurance, however, that the assumptions are valid; in fact, there is some indication in the results that some of the directional relationships in the model are misspecified. As an example, the relationships of the service use variables to grade point average are negative. Theoretically, the use of tutoring services by needy individuals should increase grade point average; a similar relationship would be expected for use of remedial services, but the theoretical argument assumes an interaction of extent of service use and need for use of that service. Need for services was not measured directly and interaction terms were not included in the structural model under consideration. The solution may reflect, therefore, that compared to individuals who do not need tutoring (or remedial) services, those who have lower grade point averages need (and therefore use) such services to a greater extent (a misspecification of directionality).

An additional limitation, indicated above, is that the procedure used is not amenable to consideration of interactions and is most interpretable under the assumption of a strictly additive model. While strict additivity of main effects is a convenient assumption, it is not one that is typically based in reality. There was some evidence in results of some preliminary analyses, not included in the report, that interactions among the variables used for prediction was, in fact, present. The assumption of a multivariate normal distribution of variables is rarely, if ever, accurate.

B. Additional Study-Specific Limitations

In general, the LISREL procedure was less than ideally suited to its application in this study. While the procedure used can be forced to accomodate unequally weighted data elements, it is not particularly suited to stratified cluster samples such as the one available in this study. In particular, the standard error estimates provided by the procedure do not incorporate the design effect reflecting stratification and clustering. The additional assumptions made in calculating analysis weights that were unrelated to sampling probabilities (see Appendix G) further restricted both the meaningfulness and generalizabilty of results. Beyond the weighting assumptions, only individuals with complete data sets were considered in these analyses, which is an implicit assumption of randomness of nonresponse. While such an assumption is convenient, it cannot be supported on theoretical or empirical grounds. As a net result of these procedures, results are applicable only to the specific sample and within that sample only to the specific subset of respondents with available data. For these reasons, results do not necessarily generalize to any well-defined larger population, and the analyses can only be viewed as exploratory.

As indicated in Section IV, some of the analytic variables used do not have the interval scale properties required by the analytic approach. Consequently, the solutions must be viewed as much more tentative than would be desirable. Moreover. the stagewise solutions described in Section IV, while required by budgetary constraints, introduce a rather large potential for errors in the results. Neither the estimated parameters of the final solution nor their associated standard errors are necessarily the estimates that would have been obtained had the solution for all parameters been obtained simultaneously.

VI. TABULAR SUMMARY OF RESULTS

Solutions of the various models examined are summarized in tabular form in this section. Resulting estimates and associated standard errors for models without measurement error (a model for receipt or nonreceipt of grants/ scholarhips and the full model) are given in Tables I.7 through I.13. Analogous results for models with estimated measurement error are given in Tables I.14 through I.22. Tables I.23, I.24, and I.26 through I.31 provide information regarding alternate computations of indirect effects. The results of tests for goodness of fit for all models examined are provided in Table I.25.

A word of caution is in order with respect to the interpretation of the test of fit of a model. In this situation, the null hypothesis is that there is no lack of fit of the model to the data. Thus, rejection of the null hypothesis as a result of a statistically significant chi-square value implies that the model does not perfectly reproduce the variance-covariance matrix which served as the input. In exploratory studies of the type reported here, this is not a serious issue. Rather, the goal is to balance the extent to which the data can be predicted from the model with the number of parameters employed. The chi-square value incorporates the number of cases in its numerator. Thus, when dealing with large sample sizes, relatively small discrepancies between the observed and predicted variance-covariance matrix tend to yield statistically significant chi-square values. In these instances, the sensitivity of the test to departure from the null hypothesis tends to eclipse the practical significance of those departures. For these reasons, the fact that all structural equation models yield predicted variance-covariance matrices that are statistically different from the observed matrices is not seen to be an important problem.

Table I.7

VALUES OF THE MATRIX Φ FOR ALL MODELING WITHOUT MEASUREMENT ERROR

	UBTYP	UBATYP	GR10	GR12	SEX	RISK	RABL	RAOTH	POVS
UBTYP	0.209								
UBATYP	-0.035	0.105							
GR10	-0.002	0.003	0.099						
GR12	0.003	-0.003	-0.054	0.250					
SEX	-0.015	-0.002	-0.004	0.016	0.239				
RISK	-0.016	-0.009	-0.001	-0.023	-0.029	0.244			
RABL	-0.003	-0.019	0.006	0.012	0.022	0.016	0.197		
RAOTH	-0.001	0.004	-0.003	-0.010	-0.016	0.001	-0.120	0.137	
POVS	-0.012	-0.021	-0.003	-0.004	0.011	0.008	0.026	-0.002	0.157

NOTE: All of the parameters in this matrix were fixed equal to their observed and/or independently estimated values. No standard errors are provided. Only the lower triangular portion of this symmetric matrix is presented.

Table I.10

SOLUTION FOR THE MATRIX Ψ FOR MODELING GRANT RECEIPT WITHOUT MEASUREMENT ERROR

	IOFF	ISIZ	ISEL	UBHOST	SSHOST	TUSE	RUSE	CUSE	GRNTR
IOFF	0.360 (0.014)								
ISIZ	0.198 (0.020)	1.389 (0.055)							
ISEL	0.181 (0.013)	0.303 (0.025)	0.524 (0.021)						
UBHOST	0.144 (0.009)	0.012 (0.016)	0.035 (0.010)	0.239 (0.009)					
SSHOST	0.091 (0.009)	0.126 (0.017)	0.006 (0.010)	0.081 (0.007)	0.246 (0.010)				
TUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.165 (0.046)			
USE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.355 (0.030)	0.923 (0.036)		
CUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.219 (0.025)	0.292 (0.023)	0.671 (0.026)	
GRNTR	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.116 (0.005)

NOTE: Cell entries are the residual variance/covariances among the endogenous variables (no parenthesis) and the standard errors of these estimates (in parentheses). Values of zero were constrained to that value on theoretical grounds. Only the lower triangular portion of this symmetric matrix is presented.

Table I.11

SOLUTION FOR THE MATRIX Γ FOR MODELING NUMBER OF TERMS,
TOTAL CREDITS EARNED AND OVERALL GPA WITHOUT MEASUREMENT ERROR

Prediction Equation For	Predictor								
	UBTYP	UBATYP	GR10	GR12	SEX	RISK	RABL	RAOTH	POVS
IOFF	0.345 (0.036)	0.172 (0.051)	0.0 (0.0)	0.0 (0.0)	-0.129 (0.028)	-0.128 (0.028)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
ISIZ	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.030 (0.103)	0.492 (0.122)	-0.285 (0.081)
ISEL	0.129 (0.043)	0.079 (0.062)	0.0 (0.0)	0.0 (0.0)	-0.146 (0.039)	-0.128 (0.039)	-0.293 (0.061)	-0.011 (0.072)	-0.090 (0.048)
UBHOST	0.202 (0.031)	0.020 (0.044)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.113 (0.031)
SSHOST	0.052 (0.031)	-0.029 (0.044)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.050 (0.026)	0.0 (0.0)	0.0 (0.0)	0.080 (0.034)
TUSE	0.206 (0.069)	0.188 (0.097)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.461 (0.100)	0.306 (0.119)	0.0 (0.0)
RUSE	0.122 (0.061)	0.106 (0.087)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.231 (0.049)	0.485 (0.090)	0.269 (0.106)	0.0 (0.0)
CUSE	0.141 (0.51)	0.156 (0.074)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.305 (0.076)	0.138 (0.091)	0.0 (0.0)
GRTCAT	0.0 (0.0)	0.0 (0.0)	0.261 (0.120)	-0.138 (0.076)	0.0 (0.0)	0.170 (0.073)	0.155 (0.120)	0.521 (0.141)	0.484 (0.091)
LOANR	0.136 (0.029)	0.093 (0.040)	-0.042 (0.042)	0.050 (0.026)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
SWORKR	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	-0.171 (0.027)	-0.080 (0.027)	-0.178 (0.043)	-0.083 (0.051)	0.0 (0.0)
NOTRM	0.0 (0.0)	0.0 (0.0)	-2.529 (0.290)	1.200 (0.183)	0.0 (0.0)	-1.485 (0.185)	1.949 (0.303)	1.006 (0.357)	-0.586 (0.233)
RTCRED	-0.679 (0.148)	-0.414 (0.205)	-1.900 (0.274)	0.824 (0.173)	0.0 (0.0)	-2.042 (0.185)	0.847 (0.307)	-0.220 (0.360)	-0.523 (0.236)
TOTGPA	-0.200 (0.046)	-0.269 (0.065)	0.0 (0.0)	0.0 (0.0)	0.117 (0.039)	-0.410 (0.043)	-0.203 (0.071)	-0.239 (0.083)	-0.121 (0.055)

NOTE: Cell entries are the coefficients (no parentheses) and standard errors of the coefficients (in parentheses). Coefficient values of zero were constrained to that value based on the result of preliminary regression solutions and/or preliminary solutions to the model. Since the model was estimated in a stagewise manner, the asymptotic standard errors were drawn from the separate estimation stages and are not, necessarily, equal to those that would be produced had the entire model been estimated simultaneously.

Table 1.12

SOLUTION FOR THE MATRIX β FOR MODELING NUMBER OF TERMS, TOTAL CREDITS EARNED AND OVERALL GPA WITHOUT MEASUREMENT ERROR

Prediction Equation for	Predictor													
	IOFF	ISIZ	ISEL	UBHOST	SSHOST	TUSE	RUSE	CUSE	GRTCAT	LOANR	SWORKR	NOTRM	RTCRED	TOTGPA
IOFF	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
ISIZ	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
ISEL	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
UBHOST	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
SSHOST	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
TUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	-0.245 (0.060)	0.147 (0.062)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
RUSE	0.0 (0.0)	0.0 (0.0)	0.127 (0.033)	0.0 (0.0)	0.180 (0.050)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
CUSE	0.0 (0.0)	0.072 (0.018)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
GRTCAT	0.0 (0.0)	0.0 (0.0)	-0.233 (0.049)	-0.398 (0.071)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
LOANR	-0.042 (0.023)	0.068 (0.011)	-0.116 (0.019)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	-0.047 (0.015)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
SWORKR	-0.092 (0.025)	0.0 (0.0)	0.0 (0.0)	0.107 (0.031)	-0.079 (0.028)	0.0 (0.0)	0.0 (0.0)	-0.060 (0.016)	0.033 (0.010)	-0.199 (0.028)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
NOTRM	-2.188 (0.147)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.995 (0.181)	-0.196 (0.058)	0.0 (0.0)	0.0 (0.0)	-0.481 (0.069)	0.0 (0.0)	-0.944 (0.184)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)
RTCRED	-2.502 (0.159)	0.334 (0.059)	-0.276 (0.095)	-0.583 (0.151)	0.605 (0.175)	0.0 (0.0)	0.0 (0.0)	0.311 (0.073)	-0.682 (0.070)	-0.412 (0.142)	-0.794 (0.187)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)
TOTGPA	0.0 (0.0)	0.100 (0.017)	0.0 (0.0)	-0.130 (0.041)	0.0 (0.0)	0.039 (0.018)	0.106 (0.020)	0.0 (0.0)	-0.052 (0.016)	-0.082 (0.044)	-0.177 (0.044)	0.0 (0.0)	0.0 (0.0)	1.00 (0.0)

NOTE: Cell entries are the coefficients (no parentheses) and standard error of the coefficients (in parentheses). Some coefficient values of in the lower triangular portion of the matrix were constrained to that value based on the result of preliminary regression solutions and/or pre solutions to the model. Diagonal elements are not part of the prediction equation. The upper triangular elements of the matrix, as well as lo triangular elements representing a crossing of variables within the same variable set (e.g., service use variables), are constrained to zero by nature of the model. Signs of the coefficients in β are the reverse of those in the structural equation model. Since the model was estimated stagewise manner, the asymptotic standard errors were drawn from the separate estimation stages and are not, necessarily, equal to those that w be produced had the entire model been estimated simultaneously.

Table 1.13

SOLUTION FOR THE MATRIX Ψ FOR MODELING NUMBER OF TERMS, TOTAL CREDITS EARNED AND OVERALL GPA WITHOUT MEASUREMENT ERROR

	IOFF	ISIZ	ISEL	UBHOST	SSHOST	TUSE	RUSE	CUSE	GRTCAT	LOANR	SWORKR	NOTRM	RTCRED	TOTGPA
IOFF	0.360 (0.014)													
ISIZ	0.198 (0.020)	1.389 (0.055)												
ISEL	0.181 (0.013)	0.303 (0.025)	0.524 (0.021)											
UBHOST	0.144 (0.009)	0.012 (0.016)	0.035 (0.010)	0.239 (0.009)										
SSHOST	0.091 (0.009)	0.126 (0.017)	0.006 (0.010)	0.081 (0.007)	0.246 (0.010)									
TUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.165 (0.046)								
RUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.355 (0.030)	0.923 (0.036)							
CUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.219 (0.025)	0.292 (0.023)	0.671 (0.026)						
GRTCAT	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.609 (0.063)					
LOANR	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.197 (0.008)				
SWORKR	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.212 (0.008)			
NOTRM	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	10.232 (0.403)		
RTCRED	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	7.403 (0.352)	10.304 (0.405)	
TOTGPA	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.693 (0.068)	1.040 (0.072)	0.544 (0.021)

NOTE: Cell entries are the residual variance/covariances among the endogenous variables (no parenthesis) and the standard errors of these estimates (in parentheses). Values of zero were constrained to that value on theoretical grounds. Only the lower triangular portion of this symmetric matrix is presented. Since the model was estimated in a stagewise manner, the asymptotic standard errors were drawn from the separate estimation stages and are not, necessarily, equal to those that would be produced had the entire model been estimated simultaneously.

Table I.14

VALUES OF THE MATRIX Φ FOR ALL MODELING
WITH ESTIMATED MEASUREMENT ERROR

	UBTYP	UBATYP	GR10	GR12	SEX	RISK	RABL	RAOTH	POVS
UBTYP	0.167								
UBATYP	-0.035	0.084							
GR10	-0.002	0.003	0.089						
GR12	0.003	-0.003	-0.054	0.225					
SEX	-0.015	-0.002	-0.004	0.016	0.227				
RISK	-0.016	-0.009	-0.001	-0.023	-0.029	0.220			
RABL	-0.003	-0.019	0.006	0.012	0.022	0.016	0.177		
RAOTH	-0.001	0.004	-0.003	-0.010	-0.016	0.001	-0.120	0.123	
POVS	-0.012	-0.021	-0.003	-0.004	0.011	0.008	0.026	-0.002	0.110

NOTE: All of the parameters in this matrix were equal to their observed and/or independently estimated values. No standard errors are provided. Only the lower triangular portion of this symmetric matrix is presented.

Table I.15

SOLUTION FOR THE MATRIX Γ
FOR MODELING GRANT RECEIPT WITH ESTIMATED MEASUREMENT ERROR

Prediction Equation for	Predictor								
	UBTYP	UBATYP	GR10	GR12	SEX	RISK	RABL	RAOTH	POVS
IOFF	0.434 (0.049)	0.270 (0.065)	0.0 (0.0)	0.0 (0.0)	-0.125 (0.031)	-0.140 (0.032)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
ISIZ	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.266 (0.174)	0.872 (0.205)	-0.434 (0.117)
ISEL	0.179 (0.057)	0.118 (0.081)	-0.168 (0.068)	-0.010 (0.043)	-0.165 (0.042)	-0.122 (0.044)	-0.473 (0.100)	-0.277 (0.117)	0.0 (0.0)
UBHOST	0.242 (0.039)	0.069 (0.055)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.152 (0.044)
SSHOST	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.062 (0.027)	0.0 (0.0)	0.228 (0.066)	0.443 (0.080)	0.0 (0.0)
TUSE	0.239 (0.086)	0.256 (0.122)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.714 (0.163)	0.592 (0.194)	0.0 (0.0)
RUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.257 (0.056)	0.759 (0.142)	0.546 (0.171)	0.0 (0.0)
CUSE	0.148 (0.063)	0.189 (0.091)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.465 (0.122)	0.319 (0.145)	0.0 (0.0)
GRNTR	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.124 (0.022)	0.220 (0.058)	0.197 (0.067)	0.163 (0.038)

NOTE: Cell entries are the coefficients (no parentheses) and standard errors of the coefficients (in parentheses). Coefficient values of zero were constrained to that value based on the result of preliminary solutions to the model. Since the model was estimated in a stagewise manner, the asymptotic standard errors were drawn from the separate estimation stages and are not, necessarily, equal to those that would be produced had the entire model been estimated simultaneously.

Table I.16

SOLUTION FOR THE MATRIX β
FOR MODELING GRANT RECEIPT WITH ESTIMATED MEASUREMENT ERROR

Prediction Equation For	Predictor								
	IOFF	ISIZ	ISEL	UBHOST	SSHOST	TUSE	RUSE	CUSE	GRNTR
IOFF	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
ISIZ	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
ISEL	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
UBHOST	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
SSHOST	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
TUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	-0.246 (0.062)	0.164 (0.064)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
RUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.199 (0.052)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)
CUSE	0.0 (0.0)	0.079 (0.022)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)
GRNTR	0.0 (0.0)	0.085 (0.014)	-0.195 (0.029)	-0.068 (0.020)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)

NOTE: Cell entries are the coefficients (no parentheses) and standard error of the coefficients (in parentheses). Some coefficient values of zero in the lower triangular portion of the matrix were constrained to that value based on the result of preliminary solutions to the model. Diagonal elements are not part of the prediction equation. The upper triangular elements of the matrix, as well as lower triangular elements representing a crossing of variables within the same variable set (e.g., service usage variables) are constrained to zero by the nature of the model. Signs of the coefficients in β are the reverse of those in the structural equation model. Since the model was estimated in a stagewise manner the asymptotic standard errors were drawn from the separate estimation stages and are not, necessarily, equal to those that would be produced had the entire model been estimated simultaneously.

Table I.17

SOLUTION FOR THE MATRIX Ψ
FOR MODELING GRANT RECEIPT WITH ESTIMATED MEASUREMENT ERROR

	IOFF	ISIZ	ISEL	UBHOST	SSHOST	TUSE	RUSE	CUSE	GRNTR
IOFF	0.292 (0.014)								
ISIZ	0.198 (0.020)	1.087 (0.055)							
ISEL	0.176 (0.013)	0.306 (0.025)	0.289 (0.020)						
UBHOST	0.141 (0.009)	0.015 (0.016)	0.032 (0.010)	0.234 (0.009)					
SSHOST	0.093 (0.009	0.120 (0.016)	0.009 (0.010)	0.083 (0.007)	0.237 (0.010)				
TUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.790 (0.046)			
RUSE	0.0 (0.0)	0 0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.350 (0.031)	0.628 (0.037)		
CUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.211 (0.025)	0.289 (0.024)	0.457 (0.026)	
GRNTR	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.071 (0.005)

NOTE: Cell entries are the residual variance/covariances among the endogenous variables (no parentheses) and the standard errors of these estimates (in parentheses). Values of zero were constrained to that value on theoretical grounds. Only the lower triangular portion of this symmetric matrix is presented. Since the model was estimated in a stagewise manner the asymptotic standard errors were drawn from the separate estimation stages and are not necessarily equal to those that would be produced had the entire model been estimated simultaneously.

Table I.18

SOLUTION FOR THE MATRICES Θ_ε AND Θ_ω

FOR MODELING GRANT RECEIPT WITH ESTIMATED MEASUREMENT ERROR

Matrix	Measurement Error for Variable Indicated								
	IOFF	ISIZ	ISES	UBHOST	SSHOST	TUSE	RUSE	CUSE	GRNTR
Θ_ε a/	0.059	0.289	0.225	0.002	0.002	0.363	0.297	0.209	0.039
	Measurement Error for Variable Indicated								
	UBTYP	UDATYP	GR10	GR12	SEX	RISK	RABL	RAOTH	POVS
Θ_ω b/	0.043 (0.007)	0.021 (0.004)	0.010 (0.003)	0.025 (0.009)	0.012 (0.009)	0.025 (0.009)	0.021 (0.004)	0.015 (0.003)	0.046 (0.006)

NOTE: Only the diagonal elements of these diagonal matrices are given.

a/ All of the parameters in this matrix were equal to their observed and/or independently estimated values. No standard errors are provided.

b/ Values in parentheses represent standard errors of the estimates. Since the model was estimated in a stagewise procedure the asymptotic standard errors were drawn from the separate estimation stages and are not, necessarily, equal to those that would be produced had the entire model been estimated simultaneously.

Table I.19

SOLUTION FOR THE MATRIX Γ FOR MODELING NUMBER OF TERMS, TOTAL CREDITS EARNED, AND OVERALL GPA WITH ESTIMATED MEASUREMENT ERROR

Prediction Equation For	Predictor								
	UBTYP	UBATYP	GR10	GR12	SEX	RISK	RABL	RAOTH	POVS
IOFF	0.434 (0.049)	0.270 (0.065)	0.0 (0.0)	0.0 (0.0)	-0.125 (0.031)	-0.140 (0.032)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
ISIZ	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.266 (0.174)	0.872 (0.205)	-0.434 (0.117)
ISEL	0.179 (0.057)	0.118 (0.081)	-0.168 (0.068)	-0.010 (0.043)	-0.165 (0.042)	-0.122 (0.044)	-0.473 (0.100)	-0.277 (0.117)	0.0 (0.0)
UBHOST	0.242 (0.039)	0.069 (0.055)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.152 (0.044)
SSHOST	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.062 (0.027)	0.0 (0.0)	0.228 (0.066)	0.443 (0.080)	0.0 (0.0)
TUSE	0.239 (0.086)	0.256 (0.122)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.714 (0.163)	0.592 (0.194)	0.0 (0.0)
RUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.257 (0.056)	0.759 (0.142)	0.546 (0.171)	0.0 (0.0)
CUSE	0.148 (0.063)	0.189 (0.091)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.465 (0.122)	0.319 (0.145)	0.0 (0.0)
GRTCAT	0.0 (0.0)	0.0 (0.0)	0.325 (0.140)	-0.137 (0.088)	0.0 (0.0)	0.193 (0.083)	0.330 (0.204)	0.722 (0.233)	0.679 (0.141)
LOANR	0.160 (0.038)	0.094 (0.052)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
SWORKR	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	-0.174 (0.029)	-0.079 (0.031)	-0.229 (0.071)	-0.155 (0.086)	0.0 (0.0)
NOTRM	0.0 (0.0)	0.0 (0.0)	-2.871 (0.356)	1.256 (0.222)	0.0 (0.0)	-1.681 (0.220)	3.318 (0.537)	2.301 (0.620)	-1.096 (0.364)
RTCRED	-1.073 (0.210)	-0.825 (0.284)	-2.231 (0.329)	0.852 (0.205)	0.331 (0.145)	-2.224 (0.230)	1.238 (0.534)	0.071 (0.621)	-1.062 (0.372)
TOTGPA	-0.300 (0.063)	-0.446 (0.090)	0.0 (0.0)	0.0 (0.0)	0.152 (0.046)	-0.422 (0.052)	-0.353 (0.124)	-0.432 (0.143)	-0.198 (0.087)

NOTE: Cell entries are the coefficients (no parentheses) and standard errors of the coefficients (in parentheses). Coefficient values of zero were constrained to that value based on the result of preliminary solutions to the model. Since the model was estimated in a stagewise manner the asymptotic standard errors were drawn from the separate estimation stages and are not, necessarily, equal to those that would be produced had the entire model been estimated simultaneously.

Table 1.20

SOLUTION FOR THE MATRIX β FOR MODELING NUMBER OF TERMS, TOTAL CREDITS EARNED, AND OVERALL GPA WITH ESTIMATED MEASUREMENT ERROR

Prediction Equation For	Predictor IOFF	ISIZ	ISEL	UBHOST	SSHOST	TUSE	RUSE	CUSE	GRTCAT	LOANR	SWORKR	NOTRM	RTCRED	TOTGPA
IOFF	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
ISIZ	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
ISEL	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
UBHOST	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
SSHOST	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
TUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	-0.248 (0.062)	0.164 (0.064)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
RUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.199 (0.052)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
CUSE	0.0 (0.0)	0.079 (0.022)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
GRTCAT	0.0 (0.0)	0.0 (0.0)	-0.401 (0.079)	-0.362 (0.074)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)	0.00 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
LOANR	0.0 (0.0)	0.111 (0.017)	-0.240 (0.034)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.079 (0.023)	-0.122 (0.029)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
SWORKR	-0.080 (0.036)	-0.039 (0.016)	0.0 (0.0)	0.081 (0.035)	-0.067 (0.030)	0.060 (0.019)	0.0 (0.0)	-0.112 (0.026)	0.049 (0.015)	-0.285 (0.043)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
NOTRM	-2.649 (0.205)	0.0 (0.0)	0.0 (0.0)	0.438 (0.169)	1.159 (0.193)	-0.223 (0.089)	0.0 (0.0)	0.0 (0.0)	-0.723 (0.110)	0.0 (0.0)	-1.410 (0.278)	1.000 (0.0)	0.0 (0.0)	0.0 (0.0)
RTCRED	-3.373 (0.192)	0.436 (0.079)	0.0 (0.0)	0.0 (0.0)	0.780 (0.183)	0.0 (0.0)	0.0 (0.0)	0.388 (0.120)	-1.052 (0.110)	-0.512 (0.217)	-1.243 (0.297)	0.0 (0.0)	1.000 (0.0)	0.0 (0.0)
TOTGPA	0.0 (0.0)	0.117 (0.022)	0.0 (0.0)	-0.114 (0.041)	0.0 (0.0)	0.0 (0.0)	0.233 (0.037)	-0.122 (0.047)	-0.093 (0.025)	0.0 (0.0)	-0.269 (0.068)	0.0 (0.0)	0.0 (0.0)	1.000 (0.0)

NOTE: Cell entries are the coefficients (no parentheses) and standard error of the coefficients (in parentheses). Some coefficient values of zero in the lower triangular portion of the matrix were constrained to that value based on the result of preliminary solutions to the model. Diagonal elements are not part of the prediction equation. The upper triangular elements of the matrix, as well as lower triangular elements representing a crossing of variables within the same variable set (e.g., service use variables) are constrained to zero by the nature of the model. Signs of the coefficients in β are the reverse of those in the structural equation model. Since the model was estimated in a stagewise manner the asymptotic standard errors were drawn from the separate estimation stages and are not, necessarily, equal to those that would be produced had the entire model been estimated simultaneously.

Table I.21

SOLUTION FOR THE MATRIX Ψ FOR MODELING
NUMBER OF TERMS, TOTAL CREDITS EARNED, AND OVERALL GPA WITH
ESTIMATED MEASUREMENT ERROR

	IOFF	ISIZ	ISEL	UBHOST	SSHOST	TUSE	RUSE	CUSE	GRTCAT	LOANR	SWORKR	NOTRM	RTCRED	TOTGPA
IOFF	0.292 (0.014)													
ISIZ	0.198 (0.020)	1.088 (0.055)												
ISEL	0.176 (0.013)	0.306 (0.025)	0.289 (0.020)											
UBHOST	0.141 (0.009)	0.014 (0.016)	0.033 (0.010)	0.234 (0.009)										
SSHOST	0.093 (0.009)	0.119 (0.016)	0.011 (0.010)	0.083 (0.007)	0.237 (0.010)									
TUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.790 (0.046)								
RUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.350 (0.031)	0.628 (0.037)							
CUSE	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.211 (0.025)	0.289 (0.024)	0.457 (0.026)						
GRTCAT	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.009 (0.064)					
LOANR	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.117 (0.008)				
SWORKR	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.125 (0.008)			
NOTRM	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	4.596 (0.419)		
RTCRED	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	6.421 (0.366)	3.947 (0.422)	
TOTGPA	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.651 (0.070)	0.946 (0.074)	0.311 (0.022)

NOTE: Cell entries are the residual variance/covariances among the endogenous variables (no parenthesis) and the standard errors of these estimates (in parentheses). Values of zero were constrained to that value on theoretical grounds. Only the lower triangular portion of this symmetric matrix is presented. Since the model was estimated in a stagewise manner the asymptotic standard errors were drawn from the separate estimation stages and are not, necessarily, equal to those that would be produced had the entire model been estimated simultaneously.

Table I.22

SOLUTIONS FOR THE MATRICES Θ_ε AND Θ_ω

FOR MODELING NUMBER OF TERMS, TOTAL CREDITS EARNED, AND OVERALL GPA WITH ESTIMATED MEASUREMENT ERROR

Matrix	Measurement Error for Variable Indicated													
	IOFF	ISIZ	ISES	UBHOST	SSHOST	TUSE	RUSE	CUSE	GRTCAT	LOANR	SWORKR	NOTRM	RTCRD	TOTGPA
Θ_ε a/	0.059	0.289	0.225	0.002	0.002	0.363	0.297	0.209	0.533	0.065	0.074	4.651	5.078	0.198
	UBTYP	UBATYP	GR10	GR12	SEX	RISK	RABL	RAOTH	POVS					
Θ_ω b/	0.043 (0.007)	0.021 (0.004)	0.010 (0.003)	0.025 (0.009)	0.012 (0.009)	0.027 (0.009)	0.022 (0.004)	0.015 (0.003)	0.047 (0.006)					

NOTE: Only the diagonal elements of these diagonal matrices are given.

a/ All of the parameters in this matrix were equal to their observed and/or independently estimated values. No standard errors are provided.

b/ Values in parentheses represent standard errors of the estimates. Since the model was estimated in a stagewise manner the asymptotic standard errors were drawn from the separate estimation stages and are not, necessarily, equal to those that would be produced had the entire model been estimated simultaneously.

Table I.23

REGRESSION SOLUTION FOR OVERALL EFFECT OF UB PARTICIPATION ON SPECIFIED OUTCOMES CONTROLLING FOR BACKGROUND FACTORS

Predictor Variable	Variable Predicted									
	TUSE	RUSE	CUSE	GRNTR	GRTCAT	LOANR	SWORKR	NOTRM	RTCRED	TOTGPA
UBTYP	.2482 (.0690)	.0985 (.0616)	.1579 (.0524)	.0771 (.0222)	.2226 (.0824)	.1720 (.0291)	.0379 (.0305)	1.0747 (.2255)	.6665 (.2412)	-.1343 (.0489)
UBATYP	.1924 (.0993)	.1117 (.0886)	.1815 (.0754)	.0260 (.0320)	.0850 (.1185)	.1116 (.0418)	.0115 (.0439)	.3599 (.3244)	.0487 (.3470)	-.2525 (.6704)

NOTE: Cell entries include (1) regression coefficients (no parentheses) for predicting specified variables, with control for other background factors and (2) standard errors for the regression coefficient (in parentheses).

Table I.24

SOLUTIONS FOR THE MATRIX Γ FOR MODELING OVERALL EFFECTS OF UB PARTICIPATION ON SPECIFIED OUTCOMES, USING MEASUREMENT ERROR AND WITH CONTROL FOR BACKGROUND FACTORS

Prediction Equation For	Predictor								
	UBTYP	UBATYP	GR10	GR12	SEX	RISK	RABL	RAOTH	POVS
TUSE	0.541 (0.091)	0.760 (0.130)	-0.249 (0.116)	-0.170 (0.073)	0.076 (0.065)	0.022 (0.071)	2.441 (0.158)	2.260 (0.181)	-0.225 (0.108)
RUSE	0.331 (0.080)	0.606 (0.114)	-0.344 (0.103)	-0.202 (0.065)	0.007 (0.058)	0.496 (0.063)	2.549 (0.145)	2.268 (0.164)	-0.410 (0.097)
CUSE	0.362 (0.069)	0.559 (0.099)	-0.145 (0.088)	0.002 (0.056)	0.064 (0.050)	0.177 (0.053)	1.587 (0.120)	1.396 (0.137)	-0.202 (0.081)
GRNTR	0.260 (0.029)	0.360 (0.042)	0.085 (0.037)	0.049 (0.023)	-0.034 (0.021)	0.100 (0.025)	-0.037 (0.056)	-0.055 (0.063)	1.131 (0.032)
GRTCAT	0.806 (0.108)	1.194 (0.154)	0.386 (0.138)	-0.076 (0.087)	-0.093 (0.078)	-0.124 (0.092)	-0.280 (0.205)	0.227 (0.231)	3.911 (0.115)
LOANR	0.254 (0.040)	0.206 (0.056)	-0.071 (0.050)	0.054 (0.032)	-0.040 (0.028)	-0.161 (0.029)	-0.081 (0.068)	-0.105 (0.077)	0.128 (0.044)
SWORKR	0.026 (0.041)	-0.035 (0.058)	-0.049 (0.052)	-0.036 (0.033)	-0.203 (0.030)	-0.279 (0.031)	-0.104 (0.071)	-0.034 (0.081)	-0.275 (0.047)
NOTRM	1.491 (0.298)	1.635 (0.422)	-3.659 (0.398)	0.673 (0.243)	-1.589 (0.213)	-7.638 (0.205)	5.662 (0.546)	4.581 (0.614)	2.511 (0.381)
RTCRED	0.962 (0.313)	1.292 (0.448)	-2.925 (0.410)	0.265 (0.253)	-1.557 (0.224)	-9.329 (0.219)	2.854 (0.595)	1.647 (0.665)	4.611 (0.405)
TOTGPA	-0.213 (0.066)	-0.338 (0.095)	-0.133 (0.085)	-0.033 (0.054)	0.014 (0.048)	-1.070 (0.049)	-0.580 (0.119)	-0.636 (0.135)	0.364 (0.079)

NOTE: Cell entries are the coefficients (no parentheses) and standard errors of the coefficients (in parentheses).

Table I.25

TESTS OF GOODNESS OF FIT OF MODELS EXAMINED

Model	CHI Square	Degrees of Freedom	Probability Level
Model for GRNTR with no error	139.97	56	0.000
Full model with no error	197.88	94[a]	0.000
Model for GRNTR with error	123.88	52[a]	0.000
Full model with error	157.10	89[a]	0.000
Model for only exogenous variables and outcomes	1,140.68	36	0.000

[a] The degrees of freedom associated with the model have been adjusted to reflect the stagewise solution of using previously obtained estimates as fixed parameter values.

Table I.26

ALTERNATE DETERMINATION OF INDIRECT EFFECTS OF UB PARTICIPATION VARIABLES
ON NUMBER OF TERMS ATTENDING
FOR MODEL WITHOUT MEASUREMENT ERROR

UB Participation	Source of Indirect Effect	Computed Indirect Effect[a/]	Standardized Total Computed Indirect Effect	Alternate Determination of Indirect Effect[b/]
Typical	Institution Type	.703	.824	(.886)
	Service Use	.049	.057	(.062)
	Financial Aid	.101	.118	(.127)
	Total	.853	1.000	1.075
Atypical	Institution Type	.376	.826	(.298)
	Service Use	.037	.081	(.029)
	Financial Aid	.042	.092	(.033)
	Total	.455	1.000	.360

a/ Determined by multiplication and addition of products of values given in Tables I.11 and I.12.

b/ Total alternate indirect effect obtained by subtracting observed direct effect (Table I.11) from alternately computed overall effect (Table I.23). Indirect effects for each listed source were determined by multiplying the total alternate indirect effect by the associated proportional effect column.

Table I.27

ALTERNATE DETERMINATION OF INDIRECT EFFECTS OF UB PARTICIPATION VARIABLES ON TERMWISE CREDITS EARNED FOR MODEL WITHOUT MEASUREMENT ERROR

UB Participation	Source of Indirect Effect	Computed Indirect Effect[a]	Standardized Total Computed Indirect Effect	Alternate Determination of Indirect Effect[b]
Typical	Institution Type	.985	.876	(1.178)
	Service Use	-.044	-.033	(-.053)
	Financial Aid	.185	.164	(.220)
	Total	1.125	1.000	1.345
Atypical	Institution Type	.430	.939	(.435)
	Service Use	-.049	-.107	(-.049)
	Financial Aid	.077	.168	(.077)
	Total	.458	1.000	.463

a/ Determined by multiplication and addition of products of values given in Tables I.11 and I.12.

b/ Total alternate indirect effect obtained by subtracting observed direct effect (Table I.11) from alternately computed overall effect (Table I.23). Indirect effects for each listed source were determined by multiplying the total alternate indirect effect by the associated proportional effect column.

Table I.28

ALTERNATE DETERMINATION OF INDIRECT EFFECTS OF UB PARTICIPATION VARIABLES ON TOTAL GPA FOR MODEL WITHOUT MEASUREMENT ERROR

UB Participation	Source of Indirect Effect	Computed Indirect Effect[a/]	Standardized Total Computed Indirect Effect	Alternate Determination of Indirect Effect[b/]
Typical	Institution Type	.026	.765	(.050)
	Service Use	-.020	-.588	(-.039)
	Financial Aid	.028	.824	(.055)
	Total	.034	1.000	.066
Atypical	Institution Type	0	.000	(0)
	Service Use	-.001	-.067	(-.001)
	Financial Aid	.016	1.067	(.018)
	Total	.015	1.000	.017

a/ Determined by multiplication and addition of products of values given in Tables I.11 and I.12.

b/ Total alternate indirect effect obtained by subtracting observed direct effect (Table I.11) from alternately computed overall effect (Table I.23). Indirect effects for each listed source were determined by multiplying the total alternate indirect effect by the associated proportional effect column.

Table I.29

ALTERNATE DETERMINATION OF INDIRECT EFFECTS OF UB PARTICIPATION VARIABLES ON NUMBER OF TERMS ATTENDING FOR MODEL WITH ESTIMATED MEASUREMENT ERROR

UB Participation	Source of Indirect Effect	Computed Indirect Effect[a/]	Standardized Total Computed Indirect Effect	Alternate Determination of Indirect Effect[b/]
Typical	Institution Type	1.044	.795	(1.185)
	Service Use	.054	.041	(.061)
	Financial Aid	.216	.164	(.245)
	Total	1.314	1.000	1.491
Atypical	Institution Type	.715	.872	(1.425)
	Service Use	.057	.070	(.114)
	Financial Aid	.048	.059	(.096)
	Total	.820	1.000	1.635

[a/] Determined by multiplication and addition of products of values given in Tables I.19 and I.20.

[b/] Total alternate indirect effect obtained by subtracting observed direct effect (Table I.19) from alternately computed overall effect (Table I.24). Indirect effects for each listed source were determined by multiplying the total alternate indirect effect by the associated proportional effect column.

Table I.30

ALTERNATE DETERMINATION OF INDIRECT EFFECTS OF UB PARTICIPATION VARIABLES ON TERMWISE CREDITS EARNED FOR MODEL WITH ESTIMATED MEASUREMENT ERROR

UB Participation	Source of Indirect Effect	Computed Indirect Effect[a/]	Standardized Total Computed Indirect Effect	Alternate Determination of Indirect Effect[b/]
Typical	Institution Type	1.464	.824	(1.677)
	Service Use	-.057	-.032	(-.064)
	Financial Aid	.370	.208	(.424)
	Total	1.777	1.000	2.035
Atypical	Institution Type	.911	1.021	(2.162)
	Service Use	-.073	-.082	(-.173)
	Financial Aid	.054	.061	(.128)
	Total	.892	1.000	2.117

a/ Determined by multiplication and addition of products of values given in Tables I.19 and I.20.

b/ Total alternate indirect effect obtained by subtracting observed direct effect (Table I.19) from alternately computed overall effect (Table I.24). Indirect effects for each listed source were determined by multiplying the total alternate indirect effect by the associated proportional effect column.

Table I.31

ALTERNATE DETERMINATION OF INDIRECT EFFECTS OF UB PARTICIPATION VARIABLES ON TOTAL GPA FOR MODEL WITH ESTIMATED MEASUREMENT ERROR

UB Participation	Source of Indirect Effect	Computed Indirect Effect[a]	Standardized Total Computed Indirect Effect	Alternate Determination of Indirect Effect[b]
Typical	Institution Type	.028	.350	(.031)
	Service Use	.018	.225	(.019)
	Financial Aid	.034	.425	(.037)
	Total	.080	1.000	.087
Atypical	Institution Type	0	.000	(0)
	Service Use	.023	.719	(.078)
	Financial Aid	.009	.281	(.030)
	Total	.032	1.000	.108

a/ Determined by multiplication and addition of products of values given in Tables I.19 and I.20.

b/ Total alternate indirect effect obtained by subtracting observed direct effect (Table I.19) from alternately computed overall effect (Table I.24). Indirect effects for each listed source were determined by multiplying the total alternate indirect effect by the associated proportional effect column.