R E P O R T   R E S U M E S

ED 015 547 EA 001 016

DROPOUT AND RETENTION RATE METHODOLOGY USED TO ESTIMATE FIRST-STAGE ELEMENTS OF THE TRANSITION PROBABILITY MATRICES FOR DYNAMOD II.

BY- ZABROWSKI, EDWARD K. HUDMAN, JOHN T.
NATIONAL CENTER FOR EDUCATIONAL STATISTICS (DHEW)

REPORT NUMBER TN-28 PUB DATE 20 APR 67
EDRS PRICE MF-$0.25 HC-$0.72 16P.

DESCRIPTORS- *DROPOUT RATE, *SCHOOL HOLDING POWER, *METHODOLOGY, ELEMENTARY SCHOOLS, SECONDARY SCHOOLS, COLLEGES, ENROLLMENT RATE, *SCHOOL DEMOGRAPHY, *PROBABILITY, STUDENTS, SEX DIFFERENCES, DISTRICT OF COLUMBIA, DYNAMOD II,

EQUATIONS FOR SYSTEM INTAKE, DROPOUT, AND RETENTION RATE CALCULATIONS ARE DERIVED FOR ELEMENTARY SCHOOLS, SECONDARY SCHOOLS, AND COLLEGES. THE PROCEDURES DESCRIBED WERE FOLLOWED IN DEVELOPING ESTIMATES OF SELECTED ELEMENTS OF THE TRANSITION PROBABILITY MATRICES USED IN DYNAMOD II. THE PROBABILITY MATRIX CELLS ESTIMATED BY THE PROCEDURES DESCRIBED ARE INCLUDED IN AN APPENDIX. A PROCEDURE FOR SEX DIFFERENTIATION IS ILLUSTRATED. (HW)

NATIONAL CENTER FOR EDUCATIONAL STATISTICS
Division of Operations Analysis

DROPOUT AND RETENTION RATE METHODOLOGY USED TO ESTIMATE FIRST-STAGE ELEMENTS OF THE TRANSITION PROBABILITY MATRICES FOR DYNAMOD II

by

Edward K. Zabrowski
John T. Hudman

Technical Note
Number 28

April 20, 1967

OFFICE OF EDUCATION/U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

NATIONAL CENTER FOR EDUCATIONAL STATISTICS
Alexander M. Mood, Assistant Commissioner

DIVISION OF OPERATIONS ANALYSIS
David S. Stoller, Director

Dropout and Retention Rate Methodology Used to Estimate First Stage Elements of the Transition Probability Matrices for DYNAMOD II

INTRODUCTION

The procedures described below were followed in the development of estimates of selected elements of the transition probability matrices used in DYNAMOD II. The probability matrix cells estimated by the procedures described in this note are designated in Appendix A.

The estimating formulas as described were used to establish first approximations of the transition matrix elements. Some of these estimates were adjusted from the formula value after comparing the model's output to the Office of Education reference data. Whenever possible, estimates were based on the 1959-60 school year. Otherwise, data were used on an "as available" basis, since this did not violate the fundamental assumption of the model--that the transition probabilities are constant for the period of calculation.

Dropout rates do not specifically appear as such in the transition matrices of DYNAMOD II, but are reflected in the larger category "other." Since dropouts constitute an important segment of the component "school leavers" (which is a critical item in the determination of retention and progression coefficients), they were estimated separately from available data for elementary and secondary school students, checked for reasonableness, and then recombined with the other elements of school leavers into an "other" transfer rate. It was felt that this would not constitute a great difficulty in later manipulations of the model, because retention rates are explicitly included in the matrices and postulated changes in dropout rates can be easily reflected in retention rates.

Two major problems were encountered in the determination of dropout and retention rates:

1. Several figures for a particular year were available, but they applied either to (a) different times in that year (Fall, June, etc.), or (b) to groups of students different from those desired. (For example, elementary school enrollment figures are not the same as those for kindergarten through grade 8, and published secondary school data are not the same as for grades 9 through 12.)1/

2. Figures for a particular year were not always available, e.g., data for odd-numbered years; and frequently the nonpublic segment of the school population was not available at the required level of detail.

METHODOLOGY

Elementary School Data

In determining the probability that an elementary school student in the Fall of 1959 would still be an elementary school student in the Fall of 1960, $P(ES_t \rightarrow ES_{t+1})$, it was assumed that the total number of elementary school students in the Fall of 1959 less dropouts, graduates, and deaths all divided by the total number of elementary school students in the Fall of 1959 would give the required probability.2/

1/ Elementary school students as defined in this paper encompass students in kindergarten through grade 8, and secondary school students are those in grades 9 through 12.

2/ Elementary school student data were obtained from the Projections of Educational Statistics to 1973-74, U.S. Office of Education, Publication Number OE-10030, 1964, p.4. Data on death rates were obtained from T. Okada, Birth and Death Projections Used in Present Student-Teacher Population Growth Models, Technical Note No. 11, December 1966.

That is:

$$(1)\quad P(ES_t \rightarrow ES_{t+1}) = \frac{ES_t - \text{"Leavers}_t\text{"}}{ES_t},$$

$$= \frac{ES_t(1-DO-DE-G_{ES})}{ES_t}, \text{ where}$$

ES_t = the number of elementary school students in year t;

DO = the dropout rate;

DE = the death rate; and

G_{ES} = the graduation rate for elementary school students who go on to secondary school

Estimates of system intake. Special complications arose in estimating elementary school dropouts, because of the discontinuity of the student flows in the lower grades. For example, a first grade student in year t+1 may or may not have attended kindergarten in year t, or he may have been a first grade repeater. Since no reliable repeater figures were available from published sources, estimates were made from two other data sources, and found to be in reasonably close agreement. The estimating formula used for repeaters was:

$$(2)\quad r_1 = \frac{n_1(8)+n_1(9)}{n_1(5)+n_1(6)+n_1(7)+n_1(8)+n_1(9)}, \text{ where}$$

r_1 = the repeater rate for the first grade; and

$n_1(i)$ = the number of first grade students in age group i.[3]

Having estimated the repeater rate, the next step was to remove its effects from the first-grade enrollment in year t+1. The balance of the first grade enrollment would then be composed of surviving kindergarten students from year t, plus those enrollees who did not attend kindergarten.[4]

$$(3)\quad (1-r_1)F_{t+1} \doteq K_t(1-DE)+(\bar{K}_t \rightarrow F_{t+1}), \text{ or}$$

$$(\bar{K}_t \rightarrow F_{t+1}) \doteq (1-r_1)F_{t+1} - K_t(1-DE), \text{ where}$$

F_{t+1} = the number of first grade students in year t+1;

K_t = the number of kindergarten students in year t; and

$(\bar{K}_t \rightarrow F_{t+1}) \doteq$ the estimate of the number of first grade enrollees in year t+1 who were not in kindergarten in year t.

Estimates of system attrition. Knowing the system's approximate intake, it was then necessary to estimate total attrition (deaths),

---

3/ Source: U.S. Department of Commerce, Bureau of the Census, U.S. Population, 1960, Summary, Detailed Characteristics, PC(1) D, Table 168. The numbers of students in ages 6 through 9 undoubtedly contain a few students who repeated kindergarten. However, it was not believed that so small a number could materially affect the estimate of the first-grade-repeater ratio.

4/ Plus a small number of students who returned to school after extended absences. Kindergarten repeaters were assumed to be too few to affect the estimate. The death rate used for kindergarteners was the same as that for the 5-14 age group. The difference in the rates between the respective groups is on the order of .0001.

(4) $G^*_{ES,t} = (n_{9,t+1} - .25c\ DO_{ss} \cdot SS_t) / 1 - \frac{DE}{4}$, where 5/

DE = the death rate for the 5-14 year age group; 6/

c = a constant value of .030 for males and .025 for females, representing the proportion of dropouts in year t reentering the system in year t+1;

$n_{9,t+1}$ = the number of ninth grade enrollments in year t+1 7/;

DO_{ss} = the dropout rate for secondary school students;

SS_t = the number of secondary school students in year t; and

$G^*_{ES,t}$ = the number of graduates from elementary school in year t.

5/ The term ".25c $DO_{ss} \cdot SS_t$" reflects information to the effect that a proportion of current secondary school dropouts will return to school t... next year. A linear estimate of the number returning to the 9th grade yields .0075 of the total male dropouts, which should be removed from the 9th grade total before calculating the transition probability. Four digits are used for computing purposes, not convey an impression of spurious accuracy. Cf. Out-of-School Youth--Two Years Later, Special Labor Force Report No. 71, August 1966, p. 861.

6/ Okada, op. cit.

7/ No estimate was possible for the small number of students (other than dropouts) who reentered the 9th grade after previously leaving the system. Note that by using the number of grade 9 enrollments to estimate the number of elementary school graduates, the graduation rate implied is only for grade 8 students who go on to the 9th grade, and does not include those students who complete elementary school but do not continue their formal education. Rough calculations were made of the number of grade 8 students who completed grade 8 but did not go to grade 9. The estimated value fell within estimated measuring error and was assumed to be statistically insignificant. Thus, $G^*_{ES,t}$ is an acceptable approximation of the overall number of graduating elementary school students.

The number of elementary students entering grade 9 in 1960 was not available, so an extrapolation was made from the nearest year for which comparable data were on hand:[8/]

(5) $$\frac{n_{9,\text{Fall 60}}}{ES_{\text{Fall 59}}} = \frac{n_{9,\text{Fall 64}}}{ES_{\text{Fall 63}}}$$

A similar result was obtained using Table 98, "Estimated Retention Rates---," on page 124 of the previously-cited <u>Digest of Educational Statistics</u>.

The graduation rate, G_{ES}, was then estimated from:

(6) $G_{ES} = \frac{G^*_{ES,t}}{ES_t}$, where $G^*_{ES,t}$ is the estimate of the number of elementary school graduates in year t.

Finally, the elementary school dropout rate, DO_{ES}, was estimated from the formula:

(7) $$DO_{ES} = \frac{ES_t(1-G_{ES}-DE) + K_{t+1} + (\bar{K}_t \longrightarrow F_{t+1}) - ES_{t+1}}{ES_t}$$

As in the case for graduates, the dropout data were keyed to 1963 and 1964 figures from the citations in footnote 7. The graduation rate in (7) was estimated for June 1964, using equations (4), (5), and (6).

The dropout figure was related to the 1959-60 school year by extrapolation:

8/ Source: U.S. Department of Health, Education, and Welfare, Office of Education, <u>Digest of Educational Statistics</u>, OE-10024-64, 1964 p. 5, and U.S. Department of Health, Education, and Welfare, Office of Education, <u>Digest of Educational Statistics</u>, OE-10024-65, 1965, page 5.

$$(8)\quad \frac{ES_{Fall\ 59}}{ES_{Fall\ 63}} = \frac{DO_{59-60}}{DO_{63-64}}$$

$$DO_{59-60} = \frac{(ES_{Fall\ 59})\ (DO_{63-64})}{ES_{Fall\ 63}}$$

The elementary school dropout rates as first calculated were not computed by sex, and were assumed to be the same for male and female students. Where the numbers of male and female dropouts were required, the proportion of male or female elementary school students was multiplied by the overall dropout rate.[9]

Secondary School Data

Estimates of system intake. The probability of an elementary student going to secondary school uses previously-determined numbers in the form:

$$(9)\quad P(ES_t \rightarrow SS_{t+1}) = (n_{9,t+1} - .25c\ DO_{ss} \cdot SS_t) / \left[(ES_t)\ (1 - \frac{DE}{4})\right]$$

where the secondary school dropout rate is as defined in (11) below.

Calculation of probabilities for secondary school (i.e., grades 9 through 12) was somewhat simpler because the number of graduates was on record.

[9] The sex distributions were taken from the listing of the Bureau of the Census' 1960 One-in-a-thousand sample data tape.

Estimates of retention and attrition. The probability that a student in secondary school will remain in secondary school the next year is given by the equation:

(10) $P(SS_t \rightarrow SS_{t+1}) = 1 - G_{ss} - (1-.75c)DO_{ss} - DE$, where

G_{ss} = the graduation rate for secondary school students; 10/

c = a constant defined in equation (4) above;

DO_{ss} = the dropout rate for secondary school students; 11/ 12/

DE = the death rate.

College Data

Estimates of system intake. The probability that a secondary school student will become a college student the next year was estimated by: 13/

10/ G_{ss} was taken to be the ratio of total high school graduates to grades 9-12 enrollment in year t. Source: Projections of Educational Statistics to 1974-75, OE-10030, op. cit., tables 2 and 14.

11/ The derivation of the secondary school student dropout rate is presented in Appendix B.

12/ Source of secondary school students by grade: see footnote 8.

13/ The data source was Ibid., pages 53 and 70. Data on the number of first-time nondegree credit enrollments could not be located, so the estimate of secondary school graduates entering degree-credit curriculums was adjusted by the ratio of degree-credit enrollments to total enrollments. Aside from error resulting from this procedures, the probability does not reflect the effect of first-year returnees.

(11) $P(SS_t \rightarrow CS_{t+1}) = G^*_{ss,t} \cdot \frac{D_{t+1}}{T_{t+1}}$, where

$G^*_{ss,t}$ = the number of secondary school graduates in year t going on to college;[14/]

D_{t+1} = total degree-program enrollees in year t+1; and

T_{t+1} = total of any-program college students in year t+1.

Distribution of intake by sex. A weighting equation was used to determine the transfer probabilities by sex:

(12) $P(SS_t \rightarrow CS_{t+1}) = w_1 ME + w_2 FE$, where

(13) $ME = a(FE)$; i.e., the male transfer rate is in some ratio to the female transfer rate, and where:

$P(SS_t \rightarrow CS_{t+1})$ = the aggregate probability that a secondary school student will become a college student the next year;

FE = the probability that a female secondary school student will enter college in the Fall;

$w_1 = \frac{\text{number of male first time college enrollees}}{\text{total number of first time college enrollees}}$; and

$w_2 = \frac{\text{number of female first time college enrollees}}{\text{total number of first time college enrollees}}$

14/ The number of secondary school students going to college was calculated by taking the number of June, 1960, graduates (from Projections of Educational Statistics, 1965 ed., OE 10030-65, table 14, p. 20) and multiplying that number by the ratio of first-time college students to high school graduates (Digest of Educational Statistics, 1965 ed., OE 10024-65, table 98, p. 124). First time college students are not necessarily degree-credit enrolled, hence the correction in equation (11).

Assuming that the college entrance probabilities are proportional to the secondary school graduations, i.e.:

$$(14) \quad a = \frac{\text{number of male secondary school graduates}}{\text{number of female secondary school graduates}}$$ 15/

The ratio from (14) was then substituted into equation (13) to obtain FE, and then rearranging (12), it is seen that:

$$(15) \quad ME = \frac{P(SS_t \rightarrow CS_{t+1}) - w_2 FE}{w_1}$$

Estimates of retention and attrition. A gross figure for the retention rate during a period is obtained from the equation:

$$(16) \quad \text{Retention Rate} = \frac{CS_t - \text{Leavers}_t}{CS_t}, \text{ where}$$

CS_t = the number of degree-credit enrolled students in year t.

Those leaving the system do so by dying, dropping out, or graduating. The "other" portion of college leavers, i.e., the proportions of graduates leaving the system and dropouts, were calculated as a residual from the difference between one and the balance of the row sum. 16/

15/ The definition of "a" established a determinate system of two equations in two unknowns (see equations (12) and (13)), and hence permitted a solution for the transfer relatives.

16/ This was possible because the transfer rates from college student status to the three teacher categories had been estimated previously. (Those procedures will be described in a separate technical note.

APPENDIX A

Transition Probability Matrix Cell Estimates Described in This Technical Note

The matrix shown in appendix table A designates with an "X" those cells for which the estimating procedures have been described in this Analytical Note. As an example of how to read the Table, refer to row 1, column 1 of the Table. The "X" entry is the estimated probability that an elementary school student in one year will be in elementary school the next year. It should be noted that the "X" entries in the Other column are residuals, representing the difference between the required row sum of one and the total of the remaining entries in that row.

The matrix presented in appendix table A describes only the probabilities of a transfer from one "occupation" to another. As the first of a two-step procedure to arrive at the final matrices used in DYNAMOD II, two such matrices were built, one for males and one for females respectively.

The population then was cycled through the two probability matrices, and initial adjustments were made to the probabilities until acceptable fits to the reference data were obtained.

To obtain the probability matrices used in the computer runs of DYNAMOD II, the two occupation matrices were combined with age transition probabilities by sex and race, resulting in four sex-race-age-occupation matrices of 164 coefficients each. The four matrices will be presented in a later report.

Appendix Table A.--Transition probability cell estimates described in this Technical Note

-Status next year-

- Status in present year -	Elementary School Student	Secondary School Student	College Student	Elementary School Teacher	Secondary School Teacher	College Teacher	Other	Dead 1/
Elem. school student	X	X					X	
Sec. school student		X	X				X	
College student			X				X	
Elem. school teacher								
Sec. school teacher								
College teacher								
Other								

1/ The procedures used for estimating death rates are described in T. Okada, Birth and Death Projections Used in Present Student-Teacher Population Growth Models, Technical Note No. 11, December 1966.

APPENDIX B

Derivation of the Secondary School Student Dropout Rate

The dropout rate for secondary school students was estimated in the following manner: Taking males as an example, consider the composition of the secondary school structure in year t to be composed of a denumerable group of students in grades 9 through 12. During the interval from years t to t+1, some of the students will graduate, some will die, some will drop out, and some of the dropouts in year t will return in year t+1. This basic composition of the block of students in t+1, given they were secondary school students in year t, is found in grades 10 through 12, assuming that the numbers of grade 9 and grade 12 repeaters are the same.

Thus, it may be asserted that the number of students in grades 10 through 12 in year t+1, ${}^{(12)}_{(10)}N_{t+1}$, can be represented by:

$${}^{(12)}_{(10)}N_{t+1} \equiv N^*(SS_t \rightarrow SS_{t+1}) + .75R^*_{t+1}, \text{ where}$$

$N^*(SS_t \rightarrow SS_{t+1})$ = the number of normally progressing secondary school students from year t to t+1, or

$= SS_t\ (1 - G_{ss} - DO_{ss} - DE)$; and

$R^*_{t+1} = c\ DO^*_{ss,t}$, i.e., the number of returning dropouts in year t+1 is "c" percent of the dropouts in year t, with "c" as defined in equation (4).

For males, as an example,

$${}^{(12)}_{(10)}N_{t+1} = SS_t\ (1 - G_{ss} - DO_{ss} - DE) + .75\ SS_t\ (.03\ DO_{ss}),$$

which, with some manipulation, is found to be

$= SS_t\ (1 - G_{ss} - .9775\ DO_{ss} - DE)$. From this relationship the male dropout rate can be solved:

$$DO_{ss} = 1.0230 \left(1 - G_{ss} - DE - \left(\frac{12}{10}\right) \frac{N_{t+1}}{SS_t}\right)$$

Note that $\left\{\frac{12}{10}\right\} N_{t+1}/SS_t$ is the implicitly-defined estimate of $P(SS_t \rightarrow SS_{t+1})$.