DOCUMENT RESUME

ED 254 148 HE 018 113

TITLE Development and Use of Error-Prone Models to Supplement Pre-Established Criteria (PEC) in Selecting Pell Grant Recipients for Validation.

INSTITUTION Advanced Technology, Inc., Reston, VA.; Westat Research, Inc., Rockville, Md.

SPONS AGENCY Office of Student Financial Assistance (ED), Washington, DC.

PUB DATE Dec 82

CONTRACT 300-80-0952

NOTE 81p.; For related documents, see HE 018 113-135 and HE 018 136-140.

PUB TYPE Statistical Data (110) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC05 Plus Postage.

DESCRIPTORS Comparative Analysis; Cost Effectiveness; Dependents; Error Patterns; *Evaluation Criteria; Family Characteristics; *Federal Aid; *Financial Aid Applicants; Higher Education; Income; *Mathematical Models; *Need Analysis (Student Financial Aid); Prediction; Predictor Variables; Self Supporting Students; Statistical Analysis; Student Characteristics; *Student Financial Aid

IDENTIFIERS Error Prone Model; *Pell Grant Program; Validation Verification and Testing Techniques

ABSTRACT

The development of a number of error-prone models to select Pell Grant recipients for validation is discussed. The 1983-1984 Pell Grant validation strategy consists of a two-stage approach: selection using Pre-Established Criteria (PEC) followed by selection using Error Prone Modeling (EPM). The database used for model development consists of a sample of 1980-1981 Pell Grant Recipients. The policy question is which students should be selected for various types of validation measures. Eight effectiveness measures are defined, and for each measure an error-prone model is developed that will identify those cases for which the corresponding type of validations will uncover the highest level of error. The data elements include: income, U.S. taxes paid, household size, nontaxable income, liquid assets, spouse income, and dependency status. The eight models are then compared in order to identify the most cost-effective approach to marginal selection for validation. The measures refer only to the payment consequences of discrepancies likely to be uncovered by the corresponding type of validation being used. Detailed appendices include EPM error tables and Automatic Interaction Detector coding categories for predictor variables. (SW)

DEVELOPMENT AND USE OF ERROR-PRONE MODELS TO SUPPLEMENT PRE-ESTABLISHED CRITERIA (PEC) IN SELECTING PELL GRANT RECIPIENTS FOR VALIDATION

Submitted to

OFFICE OF STUDENT FINANCIAL ASSISTANCE
DEPARTMENT OF EDUCATION

CONTRACT NO. 300-80-0952

DECEMBER 1982

ADVANCED TECHNOLOGY, INC.

12001 Sunrise Valley Drive
Reston, Virginia 22091

WESTAT, INC.

1650 Research Boulevard
Rockville, Maryland 20850

TABLE OF CONTENTS

LIST OF FIGURES

DEVELOPMENT AND USE OF ERROR-PRONE MODELS TO SUPPLEMENT PRE-ESTABLISHED CRITERIA (PEC) IN SELECTING PELL GRANT RECIPIENTS FOR VALIDATION

I. INTRODUCTION

A. Policy Question Addressed

The Pell Grant validation strategy for 1983-84 will possibly consist of a two-stage approach:

- Selection using Pre-Established Criteria (PEC) followed by
- Selection using Error Prone Modeling (EPM).

A further modification of validation focuses on which application items will be validated. The purpose of this paper is to report on the development of a number of error-prone models where each model aligns with a particular strategy concerning which items are to be validated. These models will be used to select cases for validation which have not already been selected using the Pre-Established Criteria. Thus the error-prone models will be used to define selections at the margin. It may be argued that selection using error-prone models should be used in place of the selection using the Pre-Established Criteria. This option can not be considered in the analysis here because of the lack of a sufficient or appropriate data base.

The data base used for model development consists of a sample of 1980-81 Pell Grant Recipients. As a result this data base can not provide information on truly non-eligible applicants selected for validation who were deterred from applying for assistance at an institution. Thus, one important component of the measure of selection effectiveness is excluded by the research methodology used for sample selection. In defense of this research strategy it should be noted that the sample was not drawn nor designed for development of error-prone models.

The policy question addressed by this research effort is, "Which cases (students) should be selected for various types of validation where type of validation is defined

by those application items which are validated?" We approach this question by first defining eight effectiveness measures. Each of these measures is specified based on one of the eight types of validation (i.e., which items would be validated). For each measure we develop an error-prone model which will identify those cases for which the corresponding type of validation will uncover the highest level of error. These eight models can then be compared in order to identify the most cost-effective approach to marginal selection for validation.

It should be emphasized that the eight criterion or effectiveness measures are not error measures as used elsewhere in the Pell Grant Quality Control Study. In other places, error was defined as the payment consequence of any discrepancies in any application or award computation item. The measures used in this research effort refer only to the payment consequences of discrepancies likely to be uncovered by the corresponding type of validation being used. For example, if we are focusing on the payment consequence of Adjusted Gross Income discrepancies, any payment consequences of discrepancies in the reporting of home value, other non-taxable income, and veteran benefits would not be included. This is an important point because the types of validation considered here are far more limited in uncovering discrepancies than the multi-faceted field protocol used in Stage One of the Quality Control study. In other words, the error likely to be removed by any model or strategy developed here will be considerably less than the error levels reported in the Stage One report.

In addition, there are other reasons why the aggregate error measures used in this work are below the total error levels reported for the Stage One effort. These reasons will be pointed out in the following discussions.

B. Nature of the Sample Data Base

The sample used to develop the models discussed later in this paper has the following limitations:

- The universe consisted only of 1980-81 Pell Grant recipients.
- Student cases selected for validation were excluded.
- Special condition filers were excluded.
- Late applicants (second semester, etc.) were excluded.
- Student and parent interviews or student record abstracts were not sufficiently complete due to non-response or missing data were excluded
- Certain 1980-81 recipients were excluded from the sample because of programmatic changes or because of assumptions or adjustments used to project the 1980-81 data base to the 1982-83 program year.

The second and sixth items listed above are among the reasons why the total error levels of this report are below the Stage One estimates.

II. CRITERION VARIABLES OR EFFECTIVENESS MEASURES

The effectiveness measures used to develop the error-prone models represent the change in payment which would likely result from discrepancies uncovered by the corresponding type of validation. Furthermore, it is assumed that once these discrepancies are uncovered they will also be corrected.

The eight validation strategies are as follows:

- Type 1 Validate Adjusted Gross Income for parents of dependent students or independent students.
- Type 2 Validate Adjusted Gross Income and U.S. taxes paid for independent students or parents of dependent students.
- Type 3 Validate Adjusted Gross Income, U.S. taxes paid, and household size for independent students and parents of dependent students and student/spouse net income for dependent students.
- Type 4 Validate Adjusted Gross Income and U.S. taxes paid for independent students and parents of dependent students and student/spouse net income for dependent students.
- Type 5 Validate Adjusted Gross Income, U.S. taxes paid, household size and non-taxable income for independent students or dependent parents and student/spouse net income for dependent students.
- Type 6 Validate Adjusted Gross Income, U.S. taxes paid, household size, non-taxable income and liquid assets for dependent parents or independent students and student/spouse net income for dependent students.

- Type 7 Validate Adjusted Gross Income, U.S. taxes paid, household size, and non-taxable income for independent students or dependent parents.

- Type 8 Validate Adjusted Gross Income and U.S. taxes paid for independent students or dependent parents and dependency status for all students.

Figure 1 summarizes the eight strategies.

The dependent variables are defined as the payment change associated with error uncovered in those items included in the associated type of validation. For example, Model 3 would be based on the payment change associated with correcting adjusted gross income, household size, and U.S. taxes paid and student/spouse net income of dependent students. Payment consequences of errors present in other application items would not be included in the dependent variable.

Practically speaking the calculation involves the difference between two calculated expected disbursements. One calculated expected payment is based on the values found in the fall Student Eligibility Report (SER) for all application items.* The second calculation uses fall SER values for all application items except for the items to be validated. For those items the best verified values uncovered during the multi-faceted field work are used. The difference between these two calculated expected payments represents the payment consequence of replacing application values with correct values for the items to be validated.

In contrast, the error measures reported in Stage One would have involved the payment consequences of replacing all fall SER values for application items with the best values uncovered by the multi-faceted field work. Thus, it is possible that because of offsets, disregards, stepped increments, interactions and interdependencies

---

*Cost of attendance and full-year enrollment status were derived from data collected in the spring visit to institutions.

Type	Data Elements for Independent Students or Dependent Parents					Dependent Students	All Students
	Adjusted Gross Income	U.S. Taxes Paid	House-hold Size	Non-Taxable Income	Liquid Assets	Student/ Spouse Net Income	Depen-dency Status
1	X						
2	X	X					
3	X	X	X			X	
4	X	X				X	
5	X	X	X	X		X	
6	X	X	X	X	X	X	
7	X	X	X	X			
8	X	X					X

FIGURE 1

VALIDATION STRATEGIES

in the Pell Grant payment formula, that errors in AGI, taxes paid, etc. do not have payment consequences in terms of our effectiveness measures when they would have payment consequences in the presence of the correct values for items which are not validated.

Finally, we have used a "net" rather than "over-award" effectiveness measure. This net error measure is assigned on a case basis. For each case, the dependent variable used in AID runs is the actual case error (student error). Cases with under-awards have negative award errors, and cases with over-awards have positive values. All error is therefore taken into consideration in the statistical calculations generated by the AID programs.

Using this net error measure the AID search technique can identify groups with large positive or negative error. The former groups will have 1) more cases with over-awards than under-awards, 2) cases with larger over-awards than under-awards, or 3) both more and larger over-awards. The same logic holds true for groups with relatively large net under-awards. Presentation of EPM results using the net error measure can take a number of forms. It is possible, for instance, to calculate the absolute error for each group identified by the model (in addition to calculating mean over-awards, under-awards, or net error for each group) simply by adding the absolute values of total under-awards to total over-awards.

We did not use the overpayment, nor underpayment, measure because it is possible that a group identified as having high levels of overpayments for those cases with overpayments might also have high levels of underpayments for other cases. Selection of this type of group would result in both overpayment and underpayment corrections without reducing the level of program expenditures.

Using the absolute value measure could also result in selection of groups which have both high overpayments and high underpayments. Furthermore, the model would not differentiate groups having high overpayments from groups with similarly high

underpayments. It is possible that use of a measure insensitive to this distinction would lead to selection of groups with only high underpayments.

III. PROJECTING THE SAMPLE DATA BASE

The sample was drawn from the universe of 1980-81 Pell Grant recipients. In order to utilize this data it was necessary to incorporate three types of adjustments or modifications to reflect the following changes between 1980-81 and 1982-83:

- Shifts in demographic composition
- Changes in economic magnitudes reflecting price level changes
- Programmatic changes in the Pell Grant program.

The first two adjustments were affected through the use of multiplicative factors. These factors were derived by comparing applicant data for the two program years.

The demographic factors were applied to the 1980-81 sample weights according to student status and age using June 1982 and June 1980 program statistics. Economic adjustment factors represent the ratio of the 1982-83 average to the 1980-81 average for the following application items for income, expense and equity values based on data from Technical Update No. 6 of OSFA's Applicant-Based Model:

- Adjusted Gross Income
- U.S. Taxes Paid
- Cash and Savings
- Dependent Student Assets
- Social Security Benefits
- Veterans Benefits
- Cost of Attendance.

The alterations made to reflect programmatic changes between 1980-81 and 1982-83 are incorporated in the formulae for the student aid index and payment

determination. This is accomplished by using the 1982-83 methodology against the 1980-81 data base. We encountered no difficulty in utilizing the 82-83 methodology with only one exception. The one exception involved student social security benefits. Revised program rules require segregating educational and non-educational benefits. However, the 1980-81 application allowed only joint reporting of current year amounts for these two benefit types. Therefore, we have had to assume that all student social security benefits were for educational purposes. This assumption will not be directly involved in the effectiveness measures since it is not a validated item in any of the eight types of validation.

The 1982-83 methodology, however, involves a significant policy shift reflected in the progressive tax rate structure. As a result, many of the highest-income recipients in the 1980-81 sample lose eligibility for the 1982-83 program year. To the extent that these middle income recipients had high error rates on adjusted gross income, their exclusion will reduce the total amount of error which can be identified by our error-prone models. This shift to a progressive tax structure is among the reasons why total error used here is below the Stage One estimates.

IV. MODEL APPROACH AND RESULTS

The discussion here reviews the preliminary results of error profiling using only selected application data as potential predictors and recipient cases which were not selected for validation. The criterion or dependent variable used in error profiling is student error as discussed above.

Data consisted of approximately 2,500 records of Basic Grant recipients who had not been selected for validation. For each recipient, the file contains data from the [redacted]tion (as recorded on computed applicant records obtained from the central [redacted]or), the student and parent questionnaires, IRS copies of income tax records, and student record abstracts. Many of these data items are used to calculate the best

verified student award which is the standard against which error is calculated. The list of potential predictors was restricted to the set of data elements available on the 1982-83 application as simulated by applying correction factors to 1980-81 applications. This was done because the original motivation for this effort was to develop new rules for selecting applicants for validation. Selection would have to be based on only the data elements actually on the application. Some application elements were eliminated a priori since they were not expected to have predictive power, leaving the following 45 potential predictors:

- Dependency status (independent or dependent student)
- Age of recipient
- Net income of the household
- Gross income of the household
- The portion of income earned by the father or independent student
- Unusual medical expenses (dollars and percentage above 20 percent of net income)
- Taxes paid by the parents or independent student
- Savings of the parents or independent student
- Net assets of dependent students
- Home value
- Home debt
- Home equity
- Value of investment assets
- Investment debt
- Net equity of investment assets
- Value of business or farm
- Business or farm debt
- Net equity in business or farm

- Net family assets
- Transaction number for the SAR
- Household size
- Number of dependents attending postsecondary institutions
- Whether or not tax figures are estimated
- Whether tax returns were assumed to have been filed
- Number of exemptions
- Adjusted gross income
- Social Security income
- Nontaxable income, other than Social Security
- Dependent student's own income
- Whether student is a citizen or eligible U.S. resident
- Student's marital status
- Student's Social Security educational benefits
- Student's Veterans educational benefits
- Student's estimated 1982-83 income
- Unreimbursed tuition
- Parent's marital status
- Value of itemized deductions for 1981
- Value of initial SAI
- Whether student lived with parents in 1981
- Whether student lived with parents in 1982
- Whether student was claimed as an exemption on parent's 1981 income tax return
- Whether student was claimed as an exemption on parent's 1982 income tax return
- Whether student received $750 in support from parents in 1981
- Whether student received $750 in support from parents in 1982.

The AID model evaluates each predictor with respect to its ability to form two separate groups very different from each other with respect to the level of error. After finding that predictor which yields this best split, the process is repeated on each of the two new groups. The process continues until one of three events occurs:

- Newly formed groups have fewer than 25 observations.
- There are over 51 groups.
- The best split does not improve prediction power enough, i.e., resulting between-group sum of squares is less than .1 percent of total sum of squares.

EPM 1 - Validating Adjusted Gross Income

The analysis for Error-Prone Model 1 (AGI only) produced 39 groups, 20 of which are final groups. These final groups are mutually exclusive and include all the cases used in the analysis. Twelve of the potential predictor variables come into play in the definition of the final groups. The largest average positive net error ($156) occurs in Group 13 while the largest average negative net error ($-60) is found in Group 14. In Group 13, there are 16 cases of over-award out of a total of 42 cases; these have a mean over-award of $443. Four cases in this group have under-awards (averaging $113), and the remaining 22 cases have no award error related to student misreporting of AGI. Two figures setting out net average error and mean over- and under-awards for EPM Model 1 are included in the figures in Appendix A of this report.

The first split in EPM 1 was imposed to separate independent from dependent students. This was done because of the fundamental differences between these two groups and because the predictor variables take on somewhat different meanings for these groups.

Independent students (Group 3) are then split on the portion of earned income earned by the student (as a portion of total student/spouse earned income). At the next level, groups are split according to whether students used tax data from a tax

return or estimated data, and on student age. Further splits were made in number of exemptions and gross income. Final splits were made on net household assets and adjusted gross income.

Dependent students (Group 2) split on the tax data source (from returns versus estimated data) at the second level. Third level splits were on number of transactions and on AGI. Subsequent splits were on taxes paid and parents' marital status. Group 17 continued to split on whether or not the student was claimed by parents as a tax deduction for the current year. Splits continued to occur for four more levels:

Sixth level:	Transaction number
Seventh level:	AGI
Eighth level:	AGI again
Ninth level:	AGI again.

The group with the highest net error, Group 13, consists of dependent students who stated that tax data (AGI, taxes, deductions, etc.) were estimated and who reported AGI's of over $25,000 on their applications. The definitions of all the groups, listed in order of net error, are found in the table for EPM 1 (below).

The order in which the predictor variables entered the AID model indicates the strength of their statistical explanatory power. This order for EPM 1 is:

- Income portion of father/student
- Source of tax data
- Number of Exemptions
- Adjusted Gross Income
- Number of Transactions
- Parent's marital status
- Claimed by Parents as Deduction '82
- Taxes Paid

- Gross Income
- Net Household Assets
- Age.

The groups formed by the AID model can be used to plot the relationship between total net error potentially removable and required number of additional validations. This estimated relationship is depicted in the graph labeled "Removable Error by Percent of Recipients" for EPM Model 1. First, the groups are ranked by size of error. Then total net error for each group is calculated by multiplying average group error by group size. Then the cumulative group sizes and total net error are calculated, expressed as percentages, and used to plot the points in the figure.

For EPM 1, we see that Group 13, (about 1.7 percent of all cases) accounts for 13.1 percent of cumulative net error. The top four groups (13, 33, 11 and 23) together account for 7.1 percent of cases and 50.1 percent of cumulative net error. Approximately 30 percent of all cases account for practically all net error associated with misreporting AGI.

EPM 2 -- Validating Adjusted Gross Income and Taxes Paid

Error-prone Model 2 (EPM 2) measures only that portion of student error attributable to error in reporting adjusted gross income (AGI) and taxes paid. The analysis for Error-prone Model 2 (AGI and taxes paid) resulted in a set of 45 groups, 23 of which are final groups. These 23 final groups are mutually exclusive and exhaustive, whereas the 22 other groups represent combinations of these 23 final groups. Fourteen of the 45 potential predictors are utilized in defining the final groups. Group 15 has the highest average error, an overpayment of $161, and Group 18 has the lowest average error, an underpayment of -$66.

The first split on EPM 2 was imposed to separate independent from dependent students. Independent students, Group 3, are then split based on portion of income

earned by the student. At the next level, independent students are split according to whether or not tax data supplied on the application was from a filed tax return, and according to age. These groups subsequently split on net income, number of exemptions, and household size. Further splits were made on net income, then net household assets, and age.

Dependent students, Group 2, were split on net income at the second level. Splits at the third level utilized taxes paid and itemized deductions. Fourth level splits were based on whether tax figures were estimated or were from tax returns. Number of transactions and marital status of parent(s) appear as fifth level split variables. Whether a student was claimed as a 1982 tax exemption, then number of transactions, then number in postsecondary education, and finally the father's income portion determine the subsequent splits. Complete definitions of the 23 final groups are presented in the figures for EPM 2 found in Appendix A of this report.

The importance of variables may be reflected by the order in which they first enter the model, as follows:

- Income portion of father/student
- Tax figures are estimated
- Net income
- Taxes paid
- Household size
- Transaction number
- Parents' marital status
- Net household assets
- Claimed as exemption in 1982
- Itemized deductions
- Age
- Number in postsecondary education
- Number of exemptions.

The graph labeled "Removable Error by Percent of Recipients" is reproduced below.

If Group 15 (about 2.3 percent of all nonvalidated students) was selected for additional validation, about 22 percent of student error attributable to AGI and taxes paid reporting could potentially be removed. Selection of groups 15, 31, 35, 33, and 41, which together account for 8.3 percent of nonvalidated students, could potentially expose 56.0 percent of cumulative net error. Since the graph depicts a decreasing slope, gains to additional validation become lower as additional students are selected. As noted earlier, this relationship is based on total student error potentially removable and thus may overstate error likely to be removed.

EPM 3 - Validating AGI, Taxes Paid, Household Size, and Student/Spouse Income

EPM 3 produced a set of forty-five groups of which 23 are final groups. Seventeen of the 45 predictors are used in defining the final groups.

Group 15 of EPM 3 has the highest net over-award error ($218); Group 36 has the highest net under-award error (-$90). Eighteen of 35 cases in Group 15 have over-awards averaging $422. Group 15 consists of independent recipients with AGI's of over $4,000 who have over 78 percent of earned income coming from the student, and who have used estimated data on 1040 items. (See tables below.)

The first split on EPM 3 is the independent-dependent split.

Independent students then split on earned income portion, then age and source of tax figures. Fourth level splits are on age and AGI, fifth level splits are on net income and number of exemptions, and sixth level splits are on AGI. The final two independent groups are determined by a split on household size.

Dependent students split on the second level based on whether or not students report living with parents in the current year. Third level splits for dependents are source of tax figures and taxes paid. Fourth level splits are again on taxes paid,

student/spouse assets, and number of transactions. Fifth level splits utilize number of exemptions, parents' marital status and number in post secondary education. The final dependent splits used three more variables: whether parents provided $750 financial assistance in the current year, net household assets, and dependent student/spouse income.

Complete group definitions are provided in the figures below.

The variables entered the AID model in the following order (after the independent/dependent split):

- Father/spouse income portion
- Lived with parents 1982
- Tax figures source
- Taxes paid
- Age
- AGI
- Transaction number
- Net income
- Number in college
- Number of exemptions
- Student assets
- Net household assets
- Parents' marital status
- Household size
- Supported by parents 1982
- Student/spouse income.

EPM 3 utilizes a more complex error measure than EPM 1 or EPM 2. Since household size error is distributed differently than income tax-related application

error the Lorenze curve of removable error by percent of recipients shows a slightly more gradual rate of vertical increase than the previous two models. Nevertheless, Group 15, with only 1.4 percent of cases accounts for 10.4 percent of cumulative net error and six groups representing 11.9 percent of the sample account for 50.0 percent of cumulative net error associated with the four data elements (AGI, taxes paid, household size, and student/spouse income).

EPM 4: Validating AGI, Taxes Paid and Student/Spouse Income

The EPM 4 error measure is drawn from verified data on dependent student income as well as family AGI and taxes paid. Data applicable to independent students is therefore identical to EPM 2.

Independent Group 17 has the highest mean overpayment of the 26 final groups ($161). Dependent Group 32 has the largest mean underpayment (-$44).

Splits on independent students are the same as those found in EPM 2, since the error matrix for independents in these two models is identical. Altogether, there were 51 groups formed by the model, 26 of which were final groups.

For dependent students, addition of error related to net dependent/spouse income produced the following results:

- Second level split on whether student lived with parents current year
- Third level splits on source of tax figures and parent marital status
- Fourth level splits on home values and source of tax figures
- Fifth level splits on taxes paid and AGI
- Sixth level splits on net income, age, home value and number of exemptions
- Seventh level splits on AGI and net household assets
- A final eighth level split on home debt.

The group definitions for EPM 4 are described fully in the figures for EPM 4 below.

Thirteen predictor variables (after status) entered the model for EPM 4, in the following order:

- Father/student income portion
- Lived with parents 1982
- Tax figure source
- Parent's marital status
- Exemptions
- Household size
- Net income
- Taxes paid
- Net household assets
- Age
- Home value
- AGI
- Home debt.

The Lorenze curve of removable error by percent of recipients shows that the bulk of error related to the data elements measured by EPM 4 are found in a small proportion of cases. Group 17 (with 2.3 percent of cases) accounts for 18.4 percent of cumulative net error. About 26 percent of this group have overpayment errors averaging $605 per recipient, while no recipients in this group had underpayments. The top five groups (17, 21, 45, 31, and 27) account for 10.8 percent of the recipient population and 56.4 percent of recipient error related to AGI, taxes paid, and dependent student/spouse net income.

EPM 5 - Validating AGI, Taxes Paid, Household Size, Nontaxable Income, and Student/Spouse Income

The EPM 5 AID model produced the maximum fifty-one groups of which 26 are final groups. The overall mean net error associated with the five data elements

evaluated in EPM 5 is more than twice as large as the error evaluated in EPM models 1, 2 and 3. (See the group summary table below.)

The highest average net error ($219) in EPM 5 is found in Group 28. Group 41 has a mean error of $208. The lowest net error figure is found in Group 26 ($-13). Almost half (16 of 40) cases in Group 29 have overpayment errors, and these overpayments average $617.

The basic independent/dependent status split was imposed on EPM 5 at the first level.

For independent students, the single second level split used the independent student income portion variable. Third level splits were on the age and AGI variables. Fourth level splits used source of tax figures and number of exemptions; fifth level splits used net household assets and net income; sixth level splits used net income again, and the final seventh level split was on age.

Dependent splits are quite complex, as can be seen in the diagram for EPM 5 below. There are 31 dependent student groups, including 16 final groups created from nine levels of group splits:

- Second level splits on AGI
- Third level splits on lived with parents and household size
- Fourth level splits on income portion of father, number of exemptions, and net household assets
- Fifth level split on source of tax figures
- Sixth level splits on AGI and gross income
- Seventh level splits on net income, cash, savings and checking value, and AGI
- Eighth level splits on income portion of father and parents' marital status

Ninth level split on home value.

Fourteen predictor variables define the 26 final groups. In addition to dependency status, the variables entering the model are:

- AGI
- Income portion of father/student
- Lived with parents 1982
- Source of tax figures
- Number of exemptions
- Gross income
- Net household assets
- Net income
- Household size
- Age
- Cash, savings, checking
- Parents' marital status
- Home value.

The Lorenze curve of potentially removable error by percent of recipients does not slope quite as steeply as Models 1-4, indicating the more complicated nature of interactions among the variables used to calculate the error measure. Nevertheless, a large proportion of error is found in the groups with high average net error. Group 29, with 1.6 percent of all cases, accounts for 7.9 percent of cumulative net error. Groups 29, 41, 25, 43 and 49 represent 8.8 percent of the population and 34.8 percent of cumulative net error.

EPM 6 - Validating AGI, Taxes Paid, Household Size, Nontaxable Income, Liquid Assets, and Student/Spouse Income

EPM 6 employs an error measure based on algebraic case error related to more Pell application items than any of the other seven models considered here. As a result, this model exposes the greatest amount of potentially removable case error. This model also produces the most complex AID output.

EPM 6 produces the maximum 51 total groups and 26 final, mutually exclusive groups. Group 39 has an average case net error of $236. Two other groups, 25 and 43, also have average net errors of over $200. Seventeen of 41 cases in Group 39 have over-awards averaging $572 per case; none in this group have under-awards.

As the diagram of AID results for EPM 6 illustrates, the splits defining the final groups are complicated.

For independents, the splits occur as follows:

- Second level splits on income portion of student
- Third level splits on AGI and age
- Fourth level splits on number of exemptions and source of tax figures
- Fifth level splits on net household assets and net income
- Sixth level splits on net income

Ten levels of group splits define the final dependent student groups:

- Second level split on AGI
- Third level splits on household size and source of tax figures
- Fourth level splits on net household assets, income portion of father and investment value
- Fifth level splits on number of exemptions and taxes paid
- Sixth level split on home value and age
- Seventh level split on number in college and net income
- Eighth level split on $750 financial assistance from parents in 1982 and income portion of father
- Ninth level split on income portion of father
- Tenth level split on income portion of father again.

The order in which the predictor variables entered the model is a rough indicator of their statistical strength in explaining group variance for this model.

- AGI

- Income portion of father/student
- Tax figures source
- Investment value
- Taxes paid
- Home value
- Net household assets
- Net income
- Household size
- Supported by parents 1982
- Age
- Number of exemptions
- Number in college.

The highest net error group, 39 consists of 1.7 percent of total cases and contributes 7.9 percent of cumulative net error. The top five groups, 39, 25, 43, 47 and 31 account for 9.9 percent of cases and 39.1 percent of error attributable to misreporting of the six sets of data elements included in the EPM 6 error measure.

EPM 7 - Validating AGI, Taxes Paid, Household Size and Nontaxable Income

EPM 7 produces the maximum 51 total groups and 26 final groups. Group 27 has the highest average net error ($238). Eighteeen of 40 cases in this group have over-awards averaging $549; two cases have under-awards averaging $125.

Independent groups are defined by seven levels of splits:

- Second level splits on income portion of student
- Third level splits on AGI and age
- Fourth level splits on source of tax figures and number of exemptions
- Fifth level splits on net household assets and net income
- Sixth level splits on household size and net income
- Seventh level splits on taxes paid and age.

Dependent students are divided into 14 final groups by the following eight sets of splits:

- Second level splits on household size
- Third level splits on taxes paid and lived with parents 1982
- Fourth level splits on parents' marital status, number of transactions, and number of exemptions
- Fifth level splits on source of tax figures
- Sixth level splits on taxes paid and AGI
- Seventh level splits on number of exemptions and household size
- Eighth level splits on income portion of father and student assets.

The predictor variables entered the AID model for EPM 7 in the following order after the imposed dependency status split:

- Income portion of student/spouse
- Household size
- AGI
- Lived with parents 1982
- Tax figures source
- Transaction number
- Net household assets
- Net income
- Taxes paid
- Number of exemptions
- Age
- Student/spouse assets
- Parents' marital status.

The Lorenze curve of removable error by percent of recipients for EPM 7 shows a relatively smooth, decreasing slope. Group 27, with 1.6 percent of cases constituted 9.3 percent of cumulative net error. The top five groups, accounting for 10.1 percent of cases, account for 43.3 percent of the total cumulative net error attributable to error in the four elements measured by EPM 7.

EPM 8 - Validating AGI, Taxes Paid and Dependency Status

EPM 8 is quite similar to EPM 2. Verification of dependency status, the extra element present in EPM 8, presents particular measurement problems for AID modeling. In those cases where Stage One analysis identified errors in dependency status, it was often difficult to establish what the students' correct awards should have been because necessary parent or student income, asset, or family data was not available. This is particularly problematic in the very small number of cases where we determined that students filing as dependents should actually have filed as independents. In addition, as stated above, a fairly large number of cases were deleted from this analysis because either 1) the cases were selected for PEC validation, or 2) the application data, "aged" to look like 1982-83 data, would not have entitled the student to an award at all. Finally, it is difficult to differentiate cases where dependency errors were "caught" independently of verification efforts linked to income verification. Cases where parents AGI is verified through tax forms may also be cases where dependency was established using tax form exemption listings. As a result, the AID model for EPM 8 failed to find any differences between students using verified and application data on dependency status. Group definitions in EPM 8 for students are therefore identical to those for dependent students in EPM 2. These are listed above in the discussion of EPM 2, and can be found in the figures for EPM 8 below. Differences in the appearance of the charts is the result of graphic artists placing the boxes differently on the diagrams.

Comparing the Eight Models: Cost-Effectiveness Issues

The ultimate purpose of estimating the eight models previously presented is to select the model and validation scheme which would be most cost-effective. Cost-effectiveness analysis is used to answer one of two related questions:

- Which alternative achieves a given or specified goal at the lowest cost; or
- Given a level of resources (costs) which alternative is most effective.

Figure 2 presents the cumulative number of cases and cumulative error potentially removable for the eleven most error-prone groups for each of the eight models. Figures 3-A and 3-B are graphic representations of this information. The table assumes that there are two million recipients not already flagged for validation or exempt from validation. Figure 3B is the leftmost lower portion of Figure 3A blown up by a factor of about four. This was done in order to allow visual separation of the line segments.

This table and accompanying graphical representation are derived from the set of eight "Average Net Student Error and Group Sizes for Final Groups" table presented in Appendix A of this report. The percent of cumulative net error and cumulative percent of cases columns from the earlier tables have been multiplied by aggregate net error and total cases based on two million students to develop the cumulative errors potentially removable and cumulative number of cases.

For example, if EPM 1 was used to select 350,000 cases for validation, about $33.6 million in error could potentially be removed while EPM 2 would potentially yield between $31.1 and $32.3 million in error removed.

In order to assess cost effectiveness, it is necessary to know the relative cost of the eight approaches. All eight require a tax form from either the parents of dependent students or the independent student. Models 4, 5, and 6, in addition, may require two (or even more) tax forms. Models 3 through 8 require additional

EPM 1		EPM 2		EPM 3		EPM 4	
Cumulative Number of Cases (1000's)	Cumulative Error Potentially Removable ($ Millions)	Cumulative Number of Cases (1000's)	Cumulative Error Potentially Removable ($ Millions)	Cumulative Number of Cases (1000's)	Cumulative Error Potentially Removable ($ Millions)	Cumulative Number of Cases (1000's)	Cumulative Error Potentially Removable ($ Millions)
34	$ 4.9	46	$ 8.1	28	$ 6.9	46	$ 8.1
62	8.6	74	12.3	70	14.5	88	13.4
112	15.5	104	15.3	118	20.8	136	17.7
142	18.8	134	18.4	170	26.1	166	20.9
204	24.6	166	21.1	200	29.4	216	24.8
242	27.5	214	25.0	238	32.7	268	28.4
282	30.0	252	27.3	292	37.3	306	30.6
350	33.6	290	29.6	330	40.7	334	32.5
408	35.8	324	31.1	366	42.9	392	35.2
444	36.4	362	32.3	448	46.2	430	37.5
538	37.8	426	34.6	492	49.0	464	39.0
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
2,000	37.5	2,000	37.6	2,000	65.5	2,000	44.0

EPM 5		EPM 6		EPM 7		EPM 8	
Cumulative Number of Cases (1000's)	Cumulative Error Potentially Removable ($ Millions)	Cumulative Number of Cases (1000's)	Cumulative Error Potentially Removable ($ Millions)	Cumulative Number of Cases (1000's)	Cumulative Error Potentially Removable ($ Millions)	Cumulative Number of Cases (1000's)	Cumulative Error Potentially Removable ($ Millions)
32	$ 6.7	34	$ 7.1	32	$ 7.3	46	$ 8.1
62	13.7	78	16.1	62	14.3	74	12.3
90	18.2	108	23.1	106	21.2	104	15.3
136	24.2	140	28.5	170	29.7	134	18.4
176	29.5	198	35.4	202	34.0	166	21.1
240	38.1	260	44.2	246	38.3	214	25.0
294	44.0	316	50.5	296	42.3	252	27.3
328	48.3	358	53.8	382	49.2	290	29.6
384	53.9	432	60.4	462	54.2	324	31.1
430	57.3	500	64.8	490	56.1	362	32.3
462	59.6	564	68.7	534	58.3	426	34.6
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
2,000	84.8	2,000	90.5	2,000	78.5	2,000	37.6

FIGURE 2

CUMULATIVE NUMBER OF CASES AND CUMULATIVE NET ERROR POTENTIALLY REMOVABLE:
ELEVEN MOST ERROR-PRONE GROUPS FOR ALTERNATIVE ERROR-PRONE MODELS

REMOVABLE ERROR USING ALTERNATE EPM'S

(IN MILLIONS OF DOLLARS)

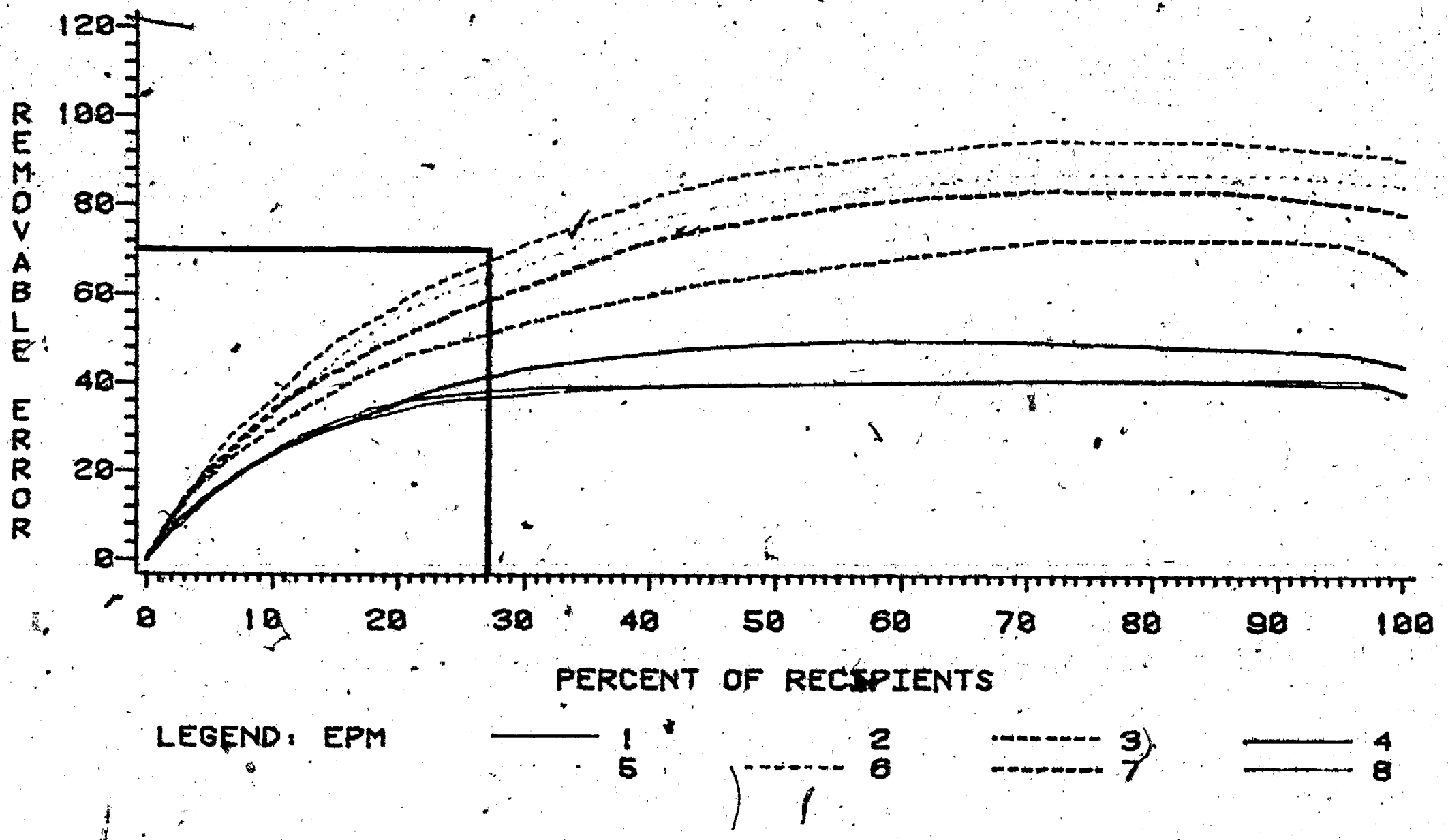


FIGURE 3-A

REMOVABLE ERROR USING ALTERNATE EPM'S

(IN MILLIONS OF DOLLARS)

(CUMULATIVE PERCENT BELOW 25%)

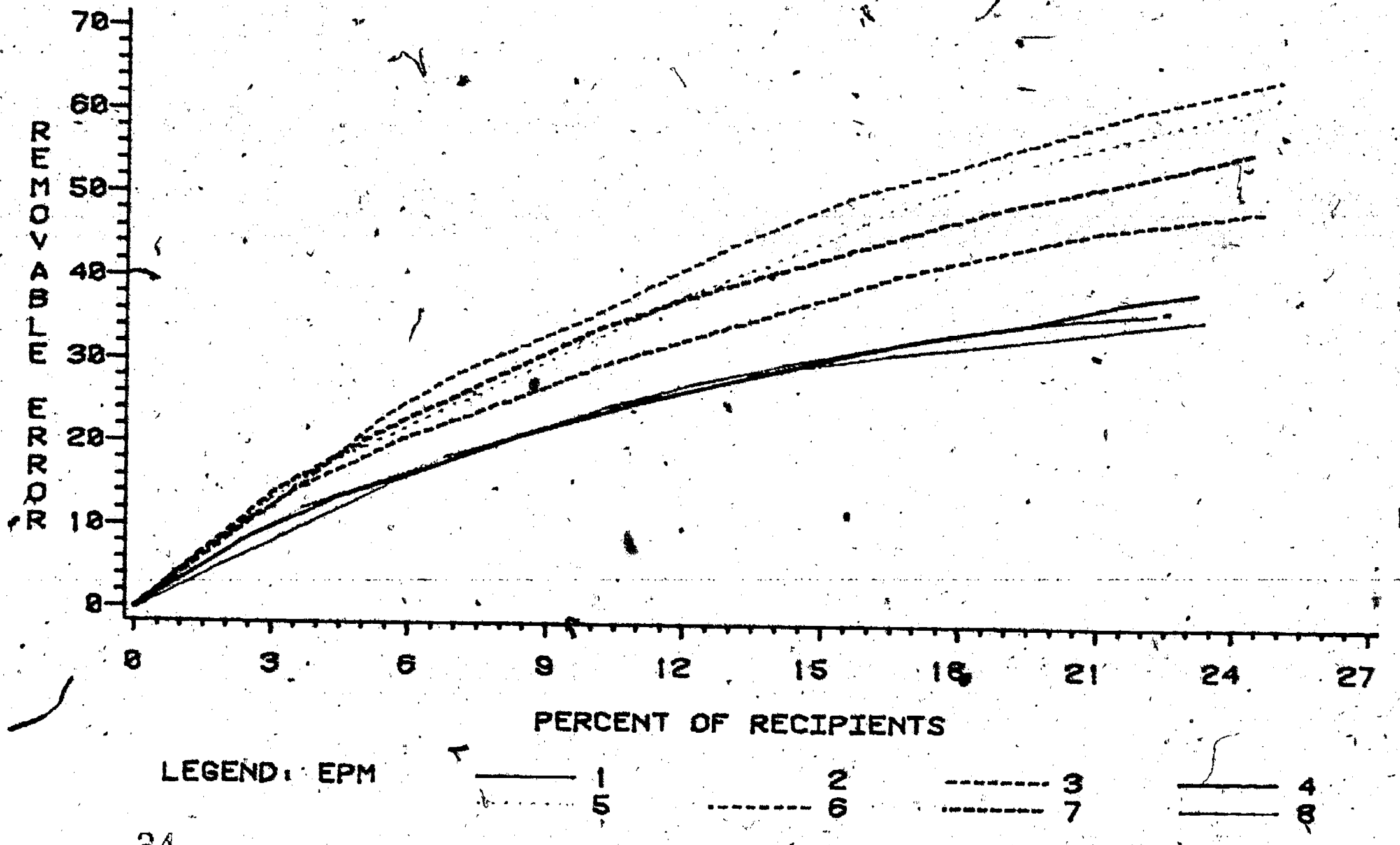


FIGURE 3-B

documentation for household size, nontaxable income, liquid assets or student status. It should also be noted that these four items are not easily documented, i.e., they would be costly to validate.

Unfortunately, we are not now able to assign relative costs to the various strategies. However, EPM 6, which could uncover about $90 million in error, is likely to be the most costly. EPM 4 is likely to be up to twice as costly as EPM 2 since EPM 4 requires documenting both student and parent income.

Removing $25 million in error would require approximately:

- 140,000 selections using EPM 5, EPM 6 or EPM 7
- 170,000 selections using EPM 3
- 204,000 selections using EPM 1
- 215,000 selections using EPM 2, EPM 4 or EPM 8.

Once the relative costs of the different schemes becomes available, or are assumed, the figures and tables can be recalculated using level of resources as the horizontal axes. This will convert the figure to represent cost-effectiveness tradeoffs amongst the eight models.

In terms of validation the first question would be, "If the Department wanted to remove $20 million of error, which scheme should be used?" As indicated below, EPM 6 would only require about 78,000 validations to achieve this objective.

Model Number	Number of Validations Required to Remove About $20 Million in Error
1	142,000
2	134,000
3	118,000
4	166,000
5	90,000
6	78,000
7	106,000
8	136,000

However, we cannot say that EPM 6 is most cost-effective unless we know or are willing to assume that cost per validation is the same for all eight models, i.e., costs will not be different. It is unlikely that costs will be the same because the number of items to be validated or documents to be collected varies across the EPM models. For example, Model 4 requires tax forms from both dependent parents and dependent students whereas Model 2 only requires one tax form per case.

In addition to the cost variation attributable to the number of required documents, the nature and/or complexity of the documents required to validate household size, liquid assets, nontaxable income and dependency status will add to the cost differences.

Therefore, the number of cases required to remove $20 million of error shown above must be converted to dollars of costs (or relative costs) in order to select the least cost method of removing the required level of error.

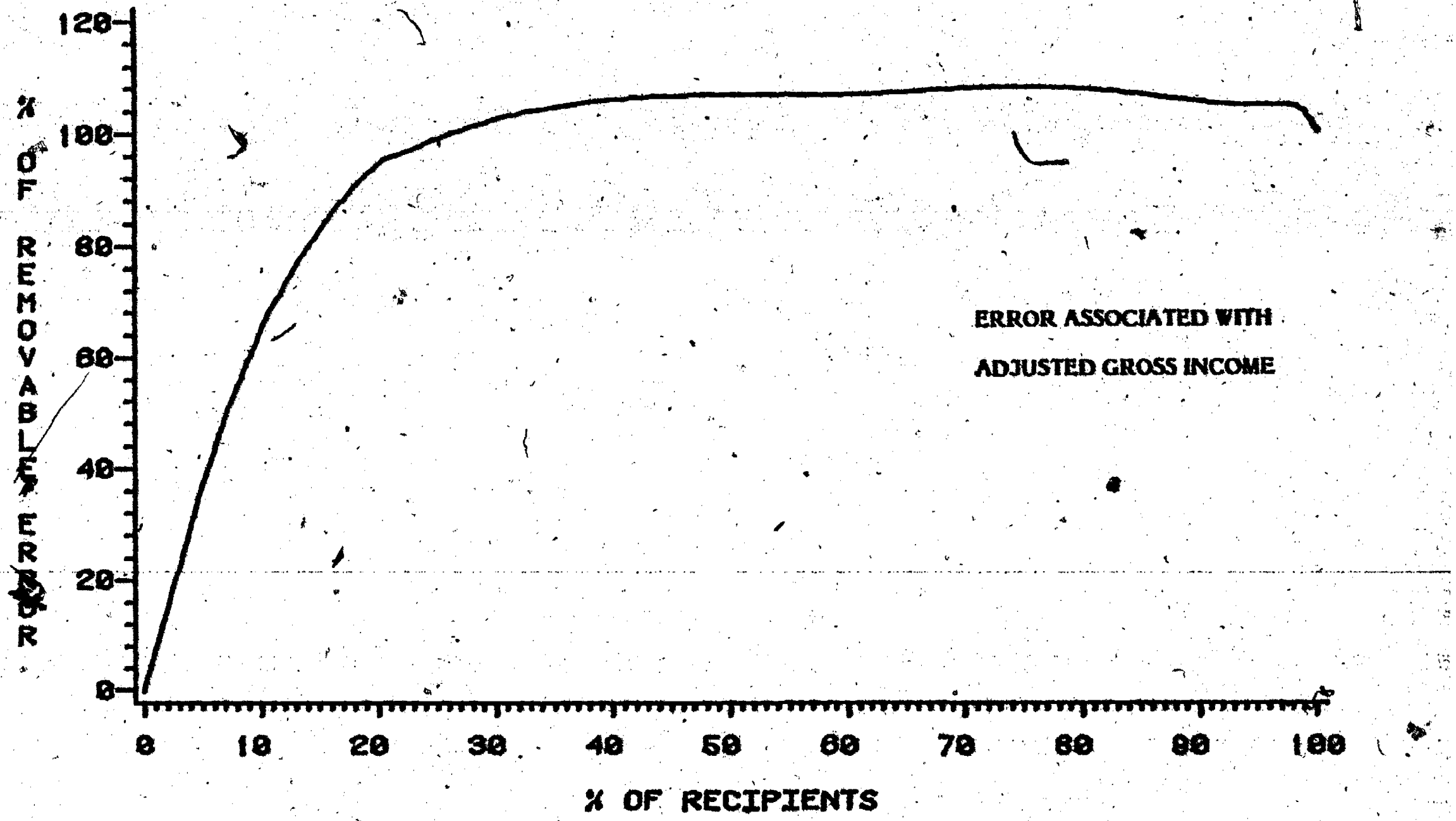
The second variant of the cost-effectiveness question requires that the number of required validations be converted to dollars of costs before we can answer the question. If, however, the costs were equal, we could ask how much error would be removed by validating up to 150,000 students:

Model Number	Error Removable by Validating Up To 150,000 Students
1	$ 18.8 Million
2	18.4
3	20.8
4	17.7
5	24.2
6	28.5
7	21.2
8	18.4

Again, it appears that EPM 6 is the most cost-effective; however, this is only because of the equal cost assumption. In reality, given the complexity of EPM 6 validation, its costs would likely be over twice that of EPM 2, EPM 1, EPM 3 and EPM 8.

REMOVABLE ERROR BY PERCENT OF RECIPIENTS

ERROR PRONE MODEL=1



GROUP NUMBER	AVERAGE NET ERROR	CUMULATIVE NET ERROR %	NUMBER OF CASES	CUMULATIVE PERCENT OF CASES
13	$156	13.1	42	1.7
33	138	22.9	35	3.1
11	126	41.4	62	5.6
23	121	50.1	37	7.1
21	87	65.6	75	10.2
37	62	73.2	48	12.1
39	61	79.9	49	14.1
29	52	89.5	85	17.5
25	41	95.3	71	20.4
20	15	96.9	44	22.2
28	14	100.8	116	26.9
18	13	103.3	101	31.0
24	2	107.4	809	63.8
26	2	108.3	167	70.5
36	1	108.5	93	74.3
30	-1	108.4	56	76.5
16	-2	107.3	179	83.8
34	-3	105.2	323	96.9
38	-4	104.9	36	98.3
14	-60	100.0	41	100.0

EPM 1:

AVERAGE NET STUDENT ERROR AND GROUP SIZES FOR FINAL GROUPS

EPM 1:

MEAN-OVERAWARDS AND UNDERAWARDS
FOR FINAL GROUPS

GROUP NUMBER	OVERAWARD MEAN	N	UNDERAWARD MEAN	N	NO ERROR N
13	$443	16	$-113	4	22
33	431	12	-49	2	21
11	988	8		0	54
23	455	11	-285	2	24
21	470	14	-12	1	60
37	495	6		0	42
39	233	15	-117	5	29
29	645	7		0	78
25	395	7		0	64
20	225	3		0	41
28	1674	1		0	115
18	173	10	-228	2	89
24	264	25	-270	16	768
26	106	5	-101	2	160
36	88	1		0	92
30	34	4	-186	1	51
16	89	5	-218	4	170
34	124	5	-341	5	313
38	238	5	-180	7	24
14	328	3	-594	6	32

Applicant Data	13	33	11	23	21	37
Status	Dependent	Dependent	Independent	Dependent	Independent	Independent
Income Portion of Father/Student			Over 78%		Over 78%	78% and Under
Source of Tax Figures	Not From Tax Form	Not From Tax Form	Not From Tax Form	Not From Tax Form	Not From Tax Form	
Number of Exemptions			0		Over 0	Over 2
Adjusted Gross Income	Over $25,000	$10,000 - $14,000		$25,000 and Under	Over $2,000	
Number of Transactions		1		2 or More		
Parent's Marital Status		Married or Divorced		Married or Divorced		
Claimed by Parents In 1982		Yes		Yes		
Taxes Paid						
Gross Income						
Net Household Assets						
Age						Over 30
Mean Group Error	$156	$139	$126	$121	$87	$62

Applicant Data	39	29	25	20	28	18
Status	Dependent	Independent	Dependent	Independent	Independent	Dependent
Income Portion of Father/Student		Over 78%		Over 78%	Over 78%	
Source of Tax Figures	Not From Tax Form	From Tax Form	From Tax Form	Not From Tax Form	From Tax Form	Not From Tax Form
Number of Exemptions				Over 0		
Adjusted Gross Income	$18,001 - $25,000		2 and Under	$2,000 and Under		$25,000 and Under
Number of Transactions	1					
Parent's Marital Status	Married or Divorced					Married or Divorced
Claimed by Parents In 1982	Yes					No
Taxes Paid			Over $3,000			
Gross Income		$5,000 and Under			$5,000 and Under	
Net Household Assets		$0 and Under			Over $0	
Age						
Mean Group Error	$61	$52	$41	$15	$14	$13

EFM 1

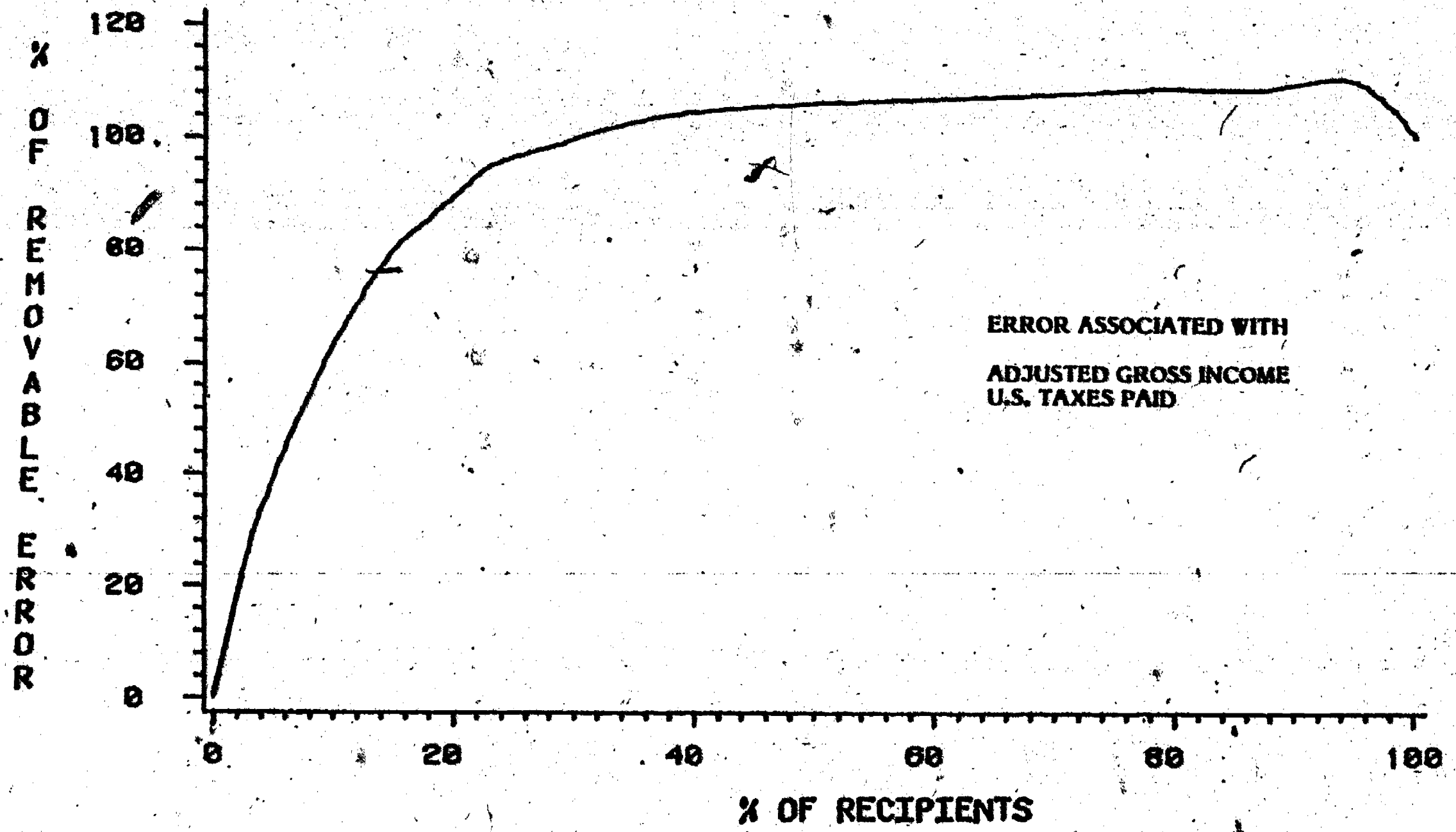
Applicant Data	24	26	36	30	16	34
Status	Dependent	Independent	Independent	Dependent	Dependent	Independent
Income Portion of Father/Student		Over 78%	78% and Under			78% and Under
Source of Tax Figures	From Tax Form	From Tax Form		Not From Tax Form	Not From Tax Form	
Number of Exemptions			2 and Under			
Adjusted Gross Income				$10,000 and Under	$25,000 and Under	
Number of Transactions	2 and Under			1		
Parent's Marital Status				Married or Divorced	Single, Widowed, and Other	
Claimed by Parents In 1982				Yes		
Taxes Paid	$3,000 and Under					
Gross Income		Over $5,000				
Net Household Assets						
Age			Over 30			30 and Under
Mean Group Error	$2	$2	$1	$-1	$-2	$-3

Applicant Data	38	14
Status	Dependent	Dependent
Income Portion of Father/Student		
Source of Tax Figures	Not From Tax Form	From Tax Form
Number of Exemptions		
Adjusted Gross Income	$14,001 - $18,000	
Number of Transactions	1	3 or More
Parent's Marital Status	Married or Divorced	
Claimed by Parents in 1982	Yes	
Taxes Paid		
Gross Income		
Net Household Assets		
Age		
Mean Group Error	$-4	$-60

EPM 1

REMOVABLE ERROR BY PERCENT OF RECIPIENTS

ERROR PRONE MODEL-2



GROUP NUMBER	AVERAGE NET ERROR	CUMULATIVE NET ERROR %	NUMBER OF CASES	CUMULATIVE PERCENT OF CASES
15	$161	21.5	57	2.3
31	160	32.6	35	3.7
35	109	40.6	38	5.2
33	94	49.0	38	6.7
41	91	56.0	40	8.3
27	75	66.5	60	10.7
11	66	72.6	46	12.6
45	50	78.7	48	14.5
32	40	82.7	43	16.2
40	34	85.9	47	18.1
26	32	91.9	80	21.3
12	29	94.9	43	23.0
30	15	96.5	53	25.1
28	13	99.1	101	29.2
38	13	101.6	96	33.1
20	7	104.3	158	39.5
24	4	104.9	70	42.3
19	2	108.4	786	74.1
44	2	108.9	93	77.9
43	1	109.4	278	89.2
22	-3	108.4	178	96.4
42	-35	104.7	45	98.2
18	-66	100.0	36	99.7

EPM 2:

AVERAGE NET STUDENT ERROR AND GROUP SIZES
FOR FINAL GROUPS

EPM 2:

MEAN OVERAWARDS AND UNDERAWARDS
FOR FINAL GROUPS

GROUP NUMBER	OVERAWARD MEAN	OVERAWARD N	UNDERAWARD MEAN	UNDERAWARD N	NO ERROR N
15	$605	15	$ 0	0	42
31	$550	10	-$ 32	1	24
35	$359	13	-$285	2	23
33	$897	4	$ 0	0	34
41	$303	13	-$ 91	5	22
27	$498	9	$ 0	0	51
11	$231	13	-$ 63	2	31
45	$373	7	-$101	2	39
32	$577	3	$ 0	0	40
40	$195	10	-$ 85	4	33
26	$631	4	-$ 12	1	75
12	$208	6	-$ 12	1	36
30	$135	12	-$385	2	39
28	$163	11	-$158	3	87
38	$165	18	-$156	10	68
20	$174	9	-$114	4	145
24	$256	1	$ 0	0	69
19	$158	50	-$187	31	705
44	$183	1	$ 0	0	92
43	$123	5	-$ 80	5	268
22	$ 53	7	-$177	5	166
42	$ 82	4	-$638	3	38
18	$304	3	-$560	6	27

	15	31	35	33	41	27
Status	Independent	Dependent	Dependent	Dependent	Dependent	Independent
Income Portion of Father/Student	Over 78%			Over 78%	Over 42%	Over 78%
Tax Figures Source	Not From Return		Not From Return	Not From Return	Not From Return	From Tax Form
Net Income	Over $2,500	Over $25,000	$25,000 and Under $3,000 and Under	$2,500 and Under	$25,000 and Under $3,000 and Under	$2,501 - $5,000
Taxes Paid						
Household Size	1		Over 1	1	1	
Transaction Number						
Parent's Marital Status			Divorced or Married		Divorced or Married	
Net Household Assets						$0
Claimed as Exemption '82			Yes		Yes	
Itemized Deductions		$1,000 and Under				
Age				22 and Under		
Number in College					1	
Exemptions						
Mean Error	161	160	109	94	91	75

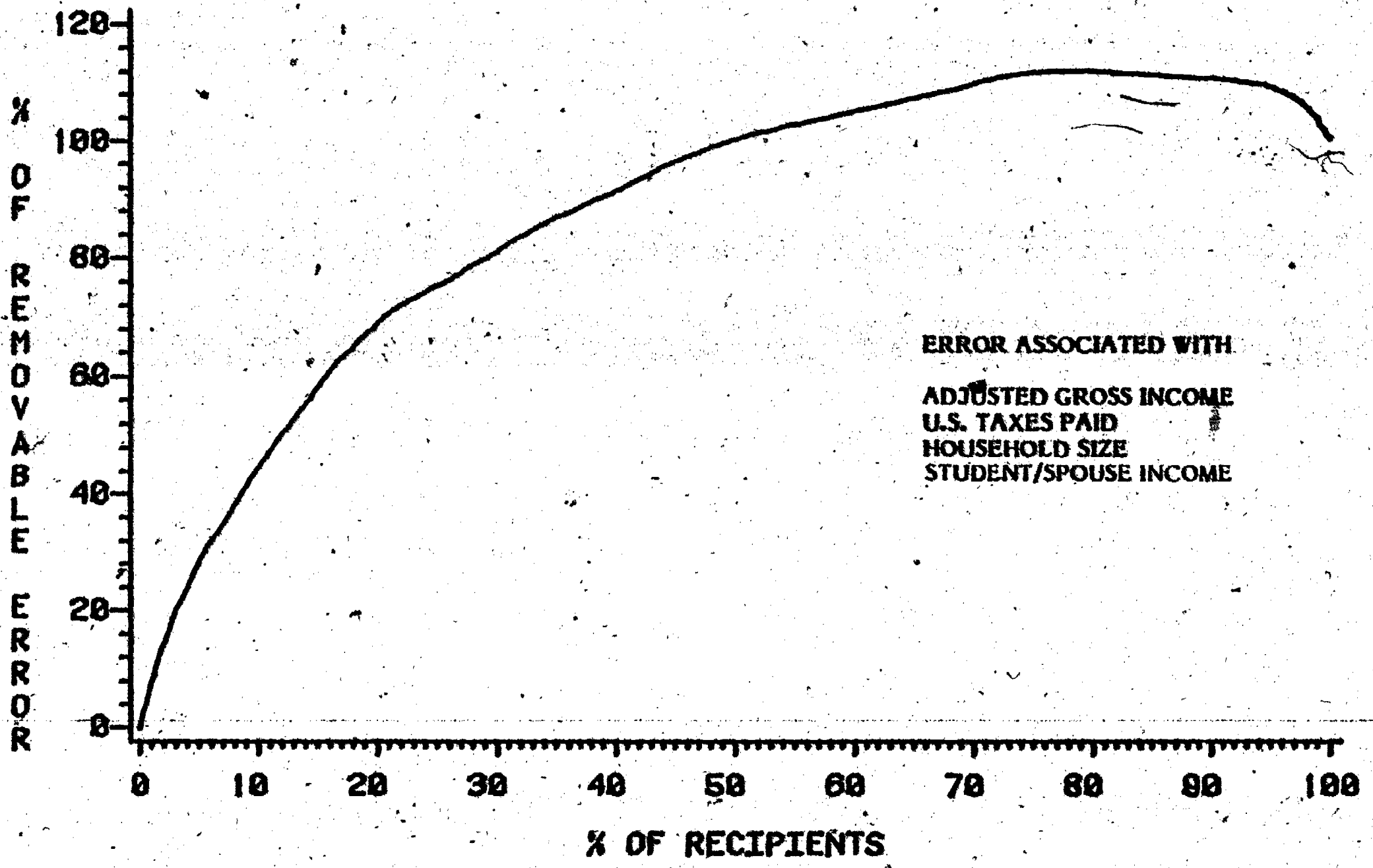
	11	45	32	40	26	12
Status	Dependent	Independent	Independent	Dependent	Independent	Independent
Income Portion of Father/Student		78% and Under	Over 78%	42% and Under	Over 78%	Over 78%
Tax Figures Source			Not From Return	Not From Return	From Tax Return	Not From Return
Net Income	$25,000 and Under Over $3,000		$2,500 and Under	$25,000 and Under $3,000 and Under	$2,501 - $5,000	
Taxes Paid						
Household Size			1	1		Over 1
Transaction Number						
Parent's Marital Status				Divorced or Married		
Net Household Assets					Over $0	
Claimed as Exemption '82				Yes		
Itemized Deductions						
Age		Over 30	Over 22			
Number in College				1		
Exemptions		Over 2				
Mean Error	66	50	40	34	32	29

	30	38	28	20	24	19
Status	Dependent	Dependent	Dependent	Independent	Independent	Dependent
Income Portion of Father/Student				Over 78%	Over 78%	
Tax Figures Source		Not From Return	Not From Return	From Tax Form	From Tax Form	From Tax Form
Net Income	Over $25,000	$25,000 and Under	$25,000 and Under	Over $5,000	$2,500 and Under	$25,000 and Under
Taxes Paid		$3,000 and Under	$3,000 and Under			$3,000 and Under
Household Size		1				2 and Under
Transaction Number						
Parent's Marital Status		Divorced or Married	Divorced or Married			
Net Household Assets						
Claimed as Exemption '82		Yes	No			
Itemized Deductions	Over $1,000					
Age						
Number in College		Over 1				
Exemptions						
Mean Error	15	13	13	7	4	2

	44	43	22	42	18
Status	Independent	Independent	Dependent	Independent	Dependent
Income Portion of Father/Student	78% and Under	78% and Under		78% and Under	
Tax Figures Source			Not From Return		From Tax Form
Net Income		$10,000 and Under	$25,000 and Under	Over $10,000	$25,000 and Under
Taxes Paid			$3,000 and Under		$3,000 and Under
Household Size					
Transaction Number					Over 2
Parent's Marital Status			Single, Widowed Separated, etc.		
Net Household Assets					
Claimed as Exemption '82					
Itemized Deductions					
Age	Over 30	30 and Under		30 and Under	
Number in College					
Exemptions	2 and Under				
Mean Error	2	1	-3	-35	-66

REMOVABLE ERROR BY PERCENT OF RECIPIENTS

ERROR PRONE MODEL-3



GROUP NUMBER	AVERAGE NET ERROR	CUMULATIVE NET ERROR %	NUMBER OF CASES	CUMULATIVE PERCENT OF CASES
15	$218	10.5	35	1.4
21	191	22.1	51	3.5
25	119	31.8	60	5.9
31	113	39.9	63	8.5
41	110	44.9	39	10.0
45	94	50.0	46	11.9
39	90	56.9	67	14.6
13	78	62.1	46	16.5
35	56	65.5	44	18.3
33	53	70.6	66	20.9
43	42	74.8	90	24.6
19	36	84.5	199	32.6
38	33	91.8	194	40.5
24	31	95.3	86	44.0
29	19	109.3	627	69.4
40	18	110.4	49	71.4
32	1	110.7	398	87.5
42	-5	110.2	96	91.4
37	-8	109.8	40	93.0
44	-9	109.4	38	94.5
28	-35	107.2	55	96.8
16	-51	104.8	41	98.4
36	-90	100.0	39	100.0

EPM 3:

AVERAGE NET STUDENT ERROR AND GROUP SIZES

FOR FINAL GROUPS

EPM 3:

MEAN OVERAWARDS AND UNDERAWARDS FOR FINAL GROUPS

GROUP NUMBER	OVERAWARD MEAN	OVERAWARD N	UNDERAWARD MEAN	UNDERAWARD N	NO ERROR N
15	$422	18	$-12	1	16
21	561	20	-213	6	25
25	1024	7		0	53
31	316	22	-58	3	38
41	317	13		0	26
45	296	17	-160	5	24
39	210	29	-183	1	37
13	663	6	-189	2	38
35	309	8		0	36
33	510	7	-93	1	58
43	203	25	-119	0	55
19	502	15	-157	3	181
38	205	48	-199	16	130
24	537	5		0	81
29	212	76	-152	25	526
40	239	6	-199	2	41
32	160	17	-253	10	371
42	100	22	-338	8	66
37	13	1	-170	2	37
44	74	2	-72	6	30
28	123	5	-242	11	39
16	198	4	-615	5	32
36	226	2	-801	5	32

Applicant Data	15	21	25	31	41	45
Status	Independent	Dependent	Independent	Dependent	Dependent	Dependent
Father/Student Income Portion	Over 78%		Over 78%			
Lived with Parents '82		No		Yes	Yes	No
Tax Figures Source	Not From Tax Form	Not From Tax Form	Not From Tax Form			From Tax Form
Taxes Paid				Over $1,000	$1,000 and Under	Over $500
Age						
Adjusted Gross Income	Over $4,000		$4,000 and Under			
Transaction Number					2 and Under	
Net Income						
Number in College					Over 2	
Number of Exemptions			0			
Student Assets				$0		
Net Household Assets						
Parent's Marital Status				Divorced		
Household Size						
Supported by Parents '82						
Student/Spouse Income					Over $0	
Mean Group Error	$218	$191	$119	$113	$110	$94

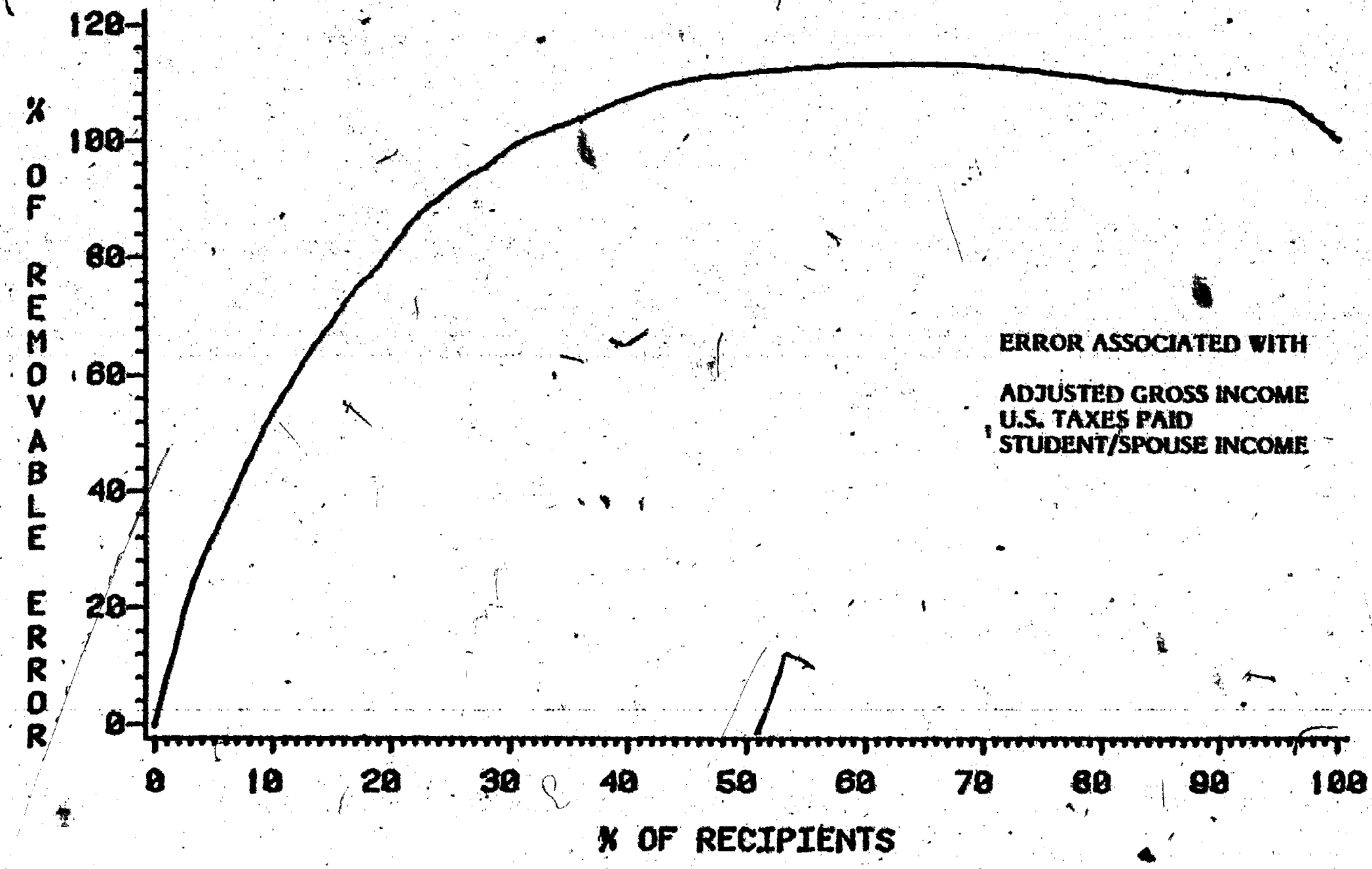
Applicant Data	39	13	35	33	43	19
Status	Dependent	Independent	Independent	Independent	Dependent	Independent
Father/Student Income Portion		Over 78%	Over 78%	78% and Under		Over 78%
Lived with Parents '82	Yes				Yes	
Tax Figures Source		From Tax Form	From Tax Form			From Tax Form
Taxes Paid	Over $1,000				Over $1,000	
Age		Over 35	35 and Under	Over 35		35 and Under
Adjusted Gross Income			Over $8,000			
Transaction Number						
Net Income			Over $5,000			$5,000 and Under
Number in College						
Number of Exemptions					Over 4	
Student Assets	$0				Over $0	
Net Household Assets						
Parent's Marital Status	Not Divorced					
Household Size						
Supported by Parents '82	No					
Student/Spouse Income						
Mean Group Error	90	78	56	54	42	36

Applicant Data	38	24	29	40	32	42
Status	Dependent	Independent	Dependent	Dependent	Independent	Dependent
Father/Student Income Portion		Over 78%			78% and Under	
Lived with Parents '82	Yes		Yes	Yes		Yes
Tax Figures Source		Not From Tax Form				
Taxes Paid	Over $1,000		$1,000 and Under	$1,000 and Under		Over $1,000
Age					35 and Under	
Adjusted Gross Income		$4,000 and Under				
Transaction Number			2 and Under	2 and Under		
Net Income						
Number in College			2 and Under	Over 2		
Number of Exemptions		Over 0				4 and Under
Student Assets	$0					Over $0
Net Household Assets			$35,000 and Under			
Parent's Marital Status	Not Divorced					
Household Size						
Supported by Parents '82	Yes					
Student/Spouse Income				$0		
Mean Group Error	$33	$31	$19	$18	$1	$-5

Applicant Data	37	44	28	16	36
Status	Independent	Dependent	Dependent	Dependent	Independent
Father/Student Income Portion	Over 78%				Over 78%
Lived with Parents '82		No	Yes	Yes	
Tax Figures Source	From Tax Form	From Tax Form			From Tax Form
Taxes Paid		$500 and Under	$1,000 and Under	$1,000 and Under	
Age	35 and Under				35 and Under
Adjusted Gross Income	$8,000 and Under				$8,000 and Under
Transaction Number			2 and Under	Over 2	
Net Income	Over $5,000				Over $5,000
Number in College			2 and Under		
Number of Exemptions					
Student Assets					
Net Household Assets			Over $35,000		
Parent's Marital Status					
Household Size	Over 1				1
Supported by Parents '82					
Student/Spouse Income					
Mean Group Error	$-8	$-9	$-35	$-51	$-90

REMOVABLE ERROR BY PERCENT OF RECIPIENTS

ERROR PRONE MODEL=4



GROUP NUMBER	AVERAGE NET ERROR	CUMULATIVE NET ERROR %	NUMBER OF CASES	CUMULATIVE PERCENT OF CASES
17	$161	18.4	57	2.3
21	133	30.4	51	4.4
45	98	40.3	60	6.8
31	94	47.4	38	8.3
27	75	56.4	60	10.8
37	69	64.6	65	13.4
39	66	69.6	46	15.3
49	65	73.8	35	16.7
43	52	80.0	71	19.6
51	50	85.3	48	21.5
30	40	88.7	43	23.2
26	32	93.8	80	26.5
44	29	95.5	35	27.9
14	29	98.1	43	29.6
47	23	103.6	146	35.6
38	19	109.3	180	42.9
22	7	111.5	158	49.3
42	5	112.2	79	52.5
24	4	112.7	70	55.3
50	2	113.1	93	59.1
40	0	113.1	205	67.4
34	-4	110.4	323	80.4
10	-5	107.4	317	93.3
48	-6	106.9	49	95.3
46	-21	105.4	46	97.0
32	-44	100.0	32	100.0

EPM 4:

AVERAGE NET STUDENT ERROR AND GROUP SIZES
FOR FINAL GROUPS

EPM 4:

MEAN OVERAWARDS AND UNDERAWARDS FOR FINAL GROUPS

GROUP NUMBER	OVERAWARD MEAN	OVERAWARD N	UNDERAWARD MEAN	UNDERAWARD N	NO ERROR N
17	$605	15	$	0	42
21	467	17	196	5	29
45	243	26	-88	5	29
31	897	4		0	34
27	498	9		0	51
37	296	15		0	50
39	277	13	-91	5	28
49	465	5	-99	2	28
43	409	10	203	1	60
51	373	7	-101	2	39
30	577	3		0	40
26	631	4	-12	1	75
44	138	11	-244	2	22
14	208	6	-12	1	36
47	129	29	-68	5	112
38	204	23	-192	7	150
22	174	9	-114	4	145
42	174	4	-144	2	73
24	256	1		0	69
50	183	1		0	92
40	59	5	-119	2	198
34	105	9	-289	8	306
10	85	12	-241	11	294
48	305	2	-187	5	42
46	111	7	-428	4	33
32	167	4	-321	12	57

Applicant Data	17	21	45	31	27	37
Status	Independent	Dependent	Dependent	Independent	Independent	Dependent
Father's/Student's Portion	Over 78%			Over 78%	Over 78%	
Lived with Parents '82		No	Yes			Yes
Tax Figures Source	Not From Return	Not From Return	Not From Return	Not From Return	From Tax Form	From Tax Form
Parents' Marital Status			Married or Divorced			Married or Divorced
Exemptions			5 and Under			
Household Size	1			1		
Net Income	Over $2,500			$2,500 and Under	$2,501 - $5,000	Over $1,500
Taxes Paid					$0 and Under	
Net Household Assets						
Age				22 or Under		Over 21
Home Value						
Adjusted Gross Income			Over $18,000			
Home Debt						
Mean Net Error	$161	$133	$98	$94	$75	$69

Applicant Data	39	49	43	51	30	26
Status	Dependent	Dependent	Dependent	Independent	Independent	Independent
Father's/Student's Portion				78% and Under	Over 78%	Over 78%
Lived with Parents '82	Yes	No	Yes			
Tax Figures Source	Not From Return	From Tax Form	From Tax Form		Not From Return	From Tax Form
Parents' Marital Status	Married or Divorced		Married or Divorced			
Exemptions				Over 2		
Household Size					1	
Net Income			$17,500 and Under		$2,500 and Under	$2,501 - $5,000
Taxes Paid			$1,500 and Under			Over $0
Net Household Assets						
Age				Over 30	Over 22	
Home Value	Over $30,000	Over $20,000				
Adjusted Gross Income	$18,000 and Under		Over $12,000			
Home Debt			$5,000 and Under			
Mean Net Error	$66	$65	$52	$50	$40	$32

EPM 4

Applicant Data	44	14	47	38	22	42
Status	Dependent	Independent	Dependent	Dependent	Independent	Dependent
Father's/Student's Portion		Over 78%			Over 78%	
Lived with Parents '82	Yes		Yes	Yes		Yes
Tax Figures Source	Not From Return	Not From Return	From Tax Form	Not From Return	From Tax Form	From Tax Form
Parents' Marital Status	Married or Divorced		Married or Divorced	Married or Divorced		Married or Divorced
Exemptions	Over 5					
Household Size		Over 1				
Taxes Paid			Over $1,500			$1,500 and Under
Net Income					Over $5,000	$17,500 and Under
Net Household Assets			Over $5,000			
Age			21 and Under			
Home Value				$30,000 and Under		
Adjusted Gross Income	Over $18,000			$18,000 and Under		Over $12,000
Home Debt						Over $5,000
Mean Net Error	$29	$29	$23	$19	$7	$5

Applicant Data	24	50	40	34	10	48
Status	Independent	Independent	Dependent	Independent	Dependent	Dependent
Father's/Student's Portion	Over 78%	78% and Under		78% and Under		
Lived with Parents '82			Yes		Yes	No
Tax Figures Source	From Tax Form		From Tax Form			From Tax Form
Parents' Marital Status			Married or Divorced		Single, Widowed and Other	
Exemptions		2 and Under				
Household Size						
Taxes Paid			$1,500 and Under			
Net Income	$2,500 and Under		$17,500 and Under			
Net Household Assets						
Age		Over 30		30 and Under		
Home Value						$20,000 and Under
Adjusted Gross Income			$12,000 and Under			
Home Debt						
Mean Net Error	$4	$2	$0	$-4	$-5	$-6

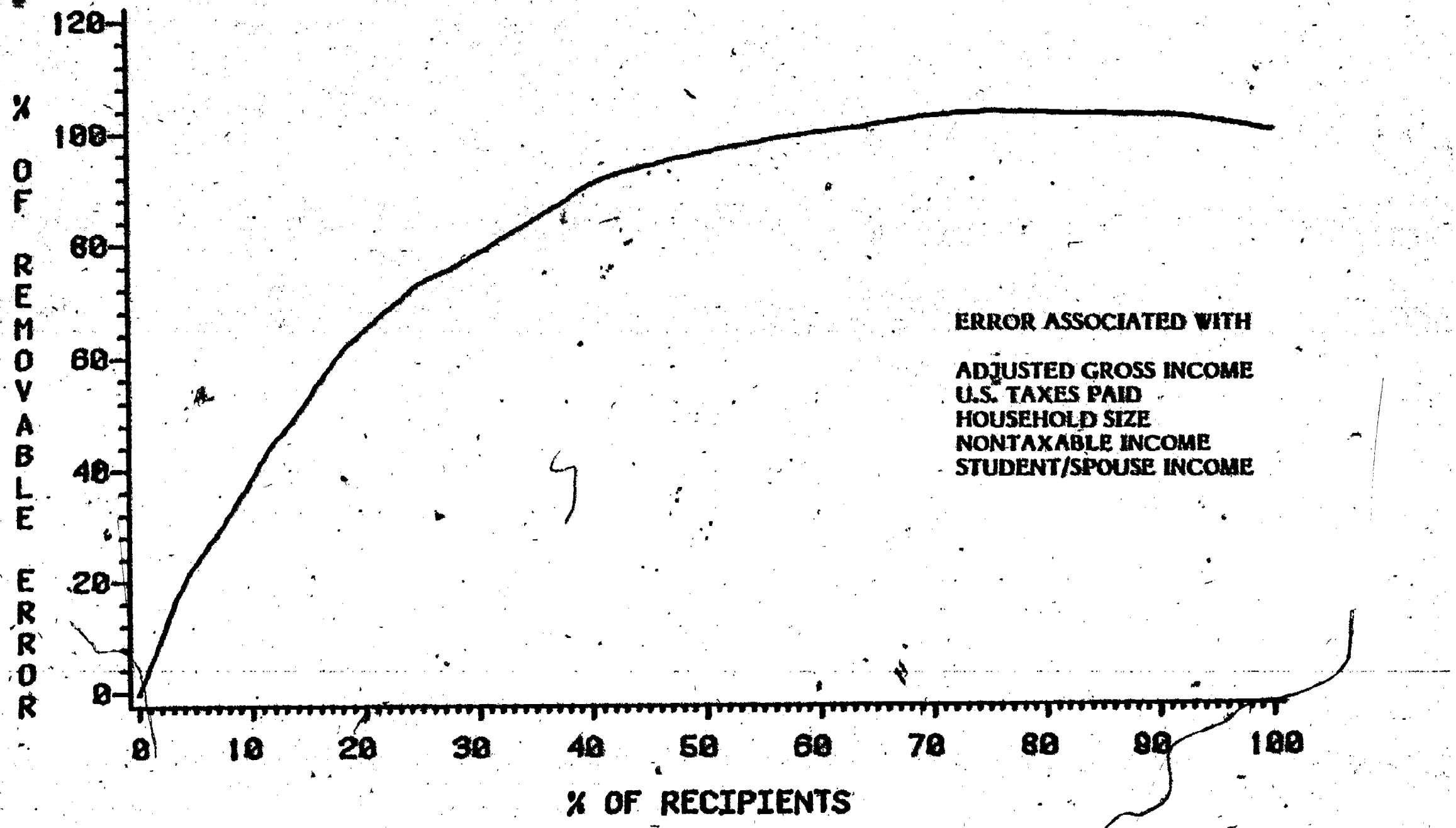
PM 4

Applicant Data	46	32
Status	Dependent	Dependent
Father's/Student's Portion		
Lived with Parents '82	Yes	Yes
Tax Figures Source	From Tax Form	From Tax Form
Parents' Marital Status	Married or Divorced	Married or Divorced
Exemptions		
Household Size		
Net Income		Over $17,500
Taxes Paid	Over $1,500	$1,500 and Under
Net Household Assets	$5,000 and Under	
Age	21 and Under	
Home Value		
Adjusted Gross Income		
Home Debt		
Mean Net Error	$-21	$-44

EPM 4

REMOVABLE ERROR BY PERCENT OF RECIPIENTS

ERROR PRONE MODEL-5



GROUP NUMBER	AVERAGE NET ERROR	CUMULATIVE NET ERROR %	NUMBER OF CASES	CUMULATIVE PERCENT OF CASES
29	$219	7.9	40	1.6
41	208	16.2	37	3.1
25	171	21.5	35	4.5
43	150	28.6	55	6.8
49	142	34.8	51	8.8
23	123	44.9	78	12.0
31	118	51.9	68	14.7
47	116	57.0	41	16.4
51	109	63.6	68	19.2
35	80	67.6	58	21.5
28	77	70.3	40	23.1
40	58	72.6	36	24.6
48	49	74.1	35	26.0
45	48	77.4	78	29.2
22	48	82.1	94	33.0
33	45	89.2	141	38.7
46	42	91.4	50	40.7
36	26	93.8	103	44.9
50	24	94.7	41	46.5
30	19	99.2	267	57.4
14	14	100.9	139	63.0
38	10	101.3	37	64.5
20	10	103.2	176	71.6
32	0	103.0	323	84.7
44	-2	102.8	145	90.6
26	-13	100.0	233	100.0

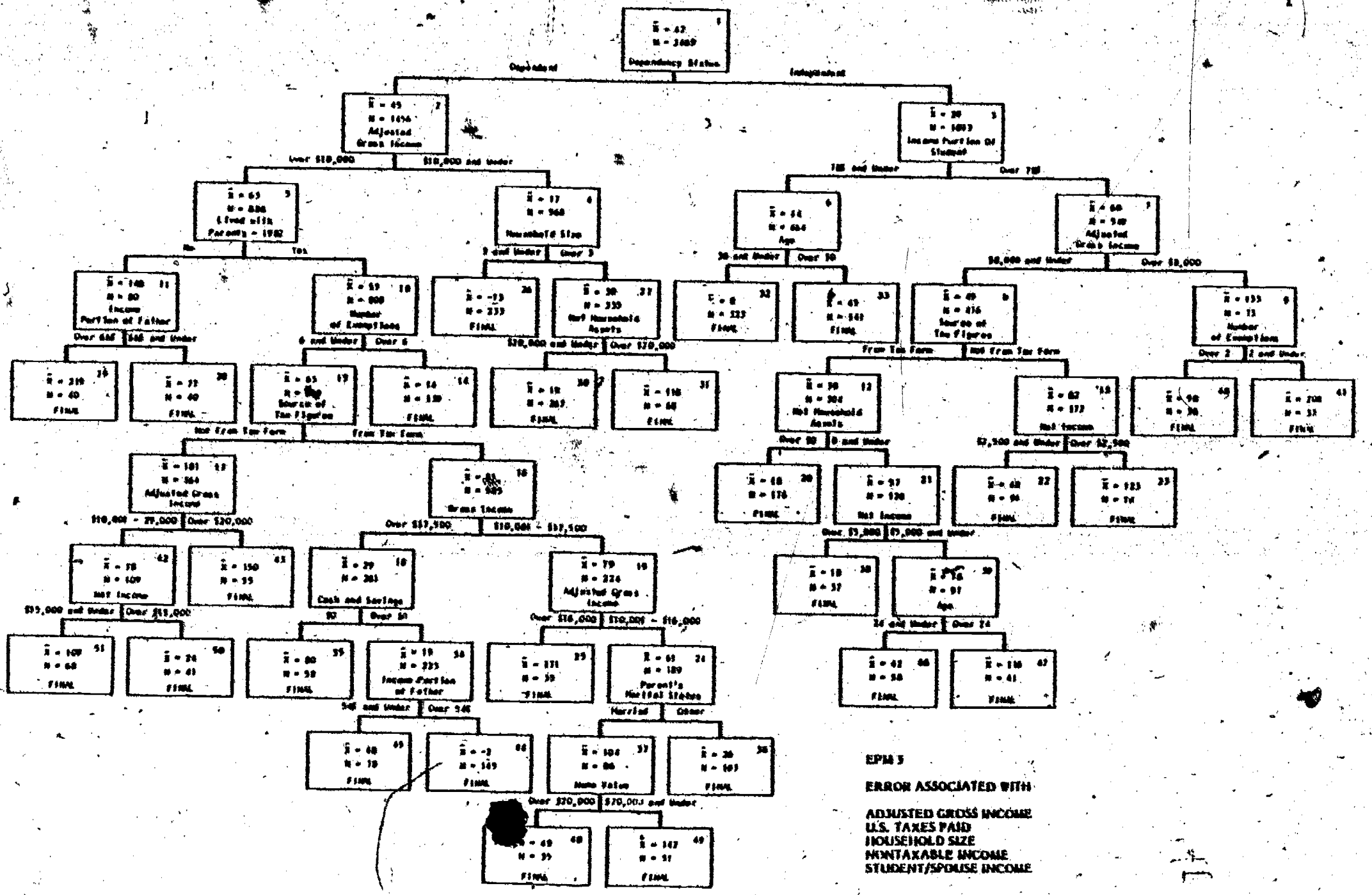
EPM 5:

AVERAGE NET STUDENT ERROR AND GROUP SIZES
FOR FINAL GROUPS

EPM 5:

MEAN OVERAWARDS AND UNDERAWARDS
FOR FINAL GROUPS

GROUP NUMBER	OVERAWARD MEAN	OVERAWARD N	UNDERAWARD MEAN	UNDERAWARD N	NO ERROR N
29	$617	16	$-202	6	18
41	518	15	- 99	1	21
25	476	13	-222	1	21
43	248	34	-81	3	18
49	522	15	-152	3	33
23	560	17		0	61
31	423	20	-182	1	47
47	687	7		0	34
51	294	27	-100	6	35
35	262	21	-207	5	32
28	232	19	-152	8	13
40	277	9	-131	3	24
48	174	12	-191	2	21
45	178	25	-135	6	47
22	570	8		0	86
33	410	16	-83	3	122
46	423	5		0	45
36	175	26	-183	10	67
50	173	15	-169	10	16
30	253	28	-198	10	229
14	170	33	-281	13	93
38	230	6	-1,004	1	30
20	486	12	-390	11	153
32	193	13	-238	11	299
44	169	38	-326	20	87
26	116	20	-183	29	184



EPM 5

ERROR ASSOCIATED WITH

ADJUSTED GROSS INCOME
U.S. TAXES PAID
HOUSEHOLD SIZE
NONTAXABLE INCOME
STUDENT/SPOUSE INCOME

BEST COPY AVAILABLE

	48	45	22	33	46	36
Status	Dependent	Dependent	Independent	Independent	Independent	Dependent
Adjusted Gross Income	$10,001 - $16,000	Over $10,000	$8,000 and Under		$8,000 and Under	$10,001 - $16,000
Income Portion of Father/Student		54% and Under	Over 78%	78% and Under	Over 78%	
Lived with Parents 1982	Yes	Yes				Yes
Source of Tax Figures	From Tax Form	From Tax Form	Not From Tax Form		From Tax Form	From Tax Form
Number of Exemptions	6 and Under	6 and Under				6 and Under
Gross Income	$10,001 - $17,500	Over $17,500				$10,001 - $17,500
Net Household Assets					$0 and Under	
Net Income			$2,500 and Under		$5,000 and Under	
Household Size						
Age				Over 30	24 and Under	
Cash and Savings		Over $0				
Parent's Marital Status	Married					Not Married
Home Value	Over $20,000					
Mean Group Error	49	48	48	45	42	26

	50	30	14	38	20	32
Status	Dependent	Dependent	Dependent	Independent	Independent	Independent
Adjusted Gross Income	$10,001 - $20,000	$10,000 and Under	Over $10,000	$8,000 and Under	$8,000 and Under	
Income Portion of Father/Student				Over 78%	Over 78%	78% and Under
Lived with Parents 1982	Yes		Yes			
Source of Tax Figures	Not From Tax Form			From Tax Form	From Tax Form	
Number of Exemptions	6 and Under		Over 6			
Gross Income		$20,000 and Under		$0 and Under		
Net Household Assets	Over $15,000			Over $5,000	Over $0	
Net Income						
Household Size		Over 3				
Age						30 and Under
Cash and Savings						
Parent's Marital Status						
Home Value						
Mean Group Error	24	19	14	10	10	0

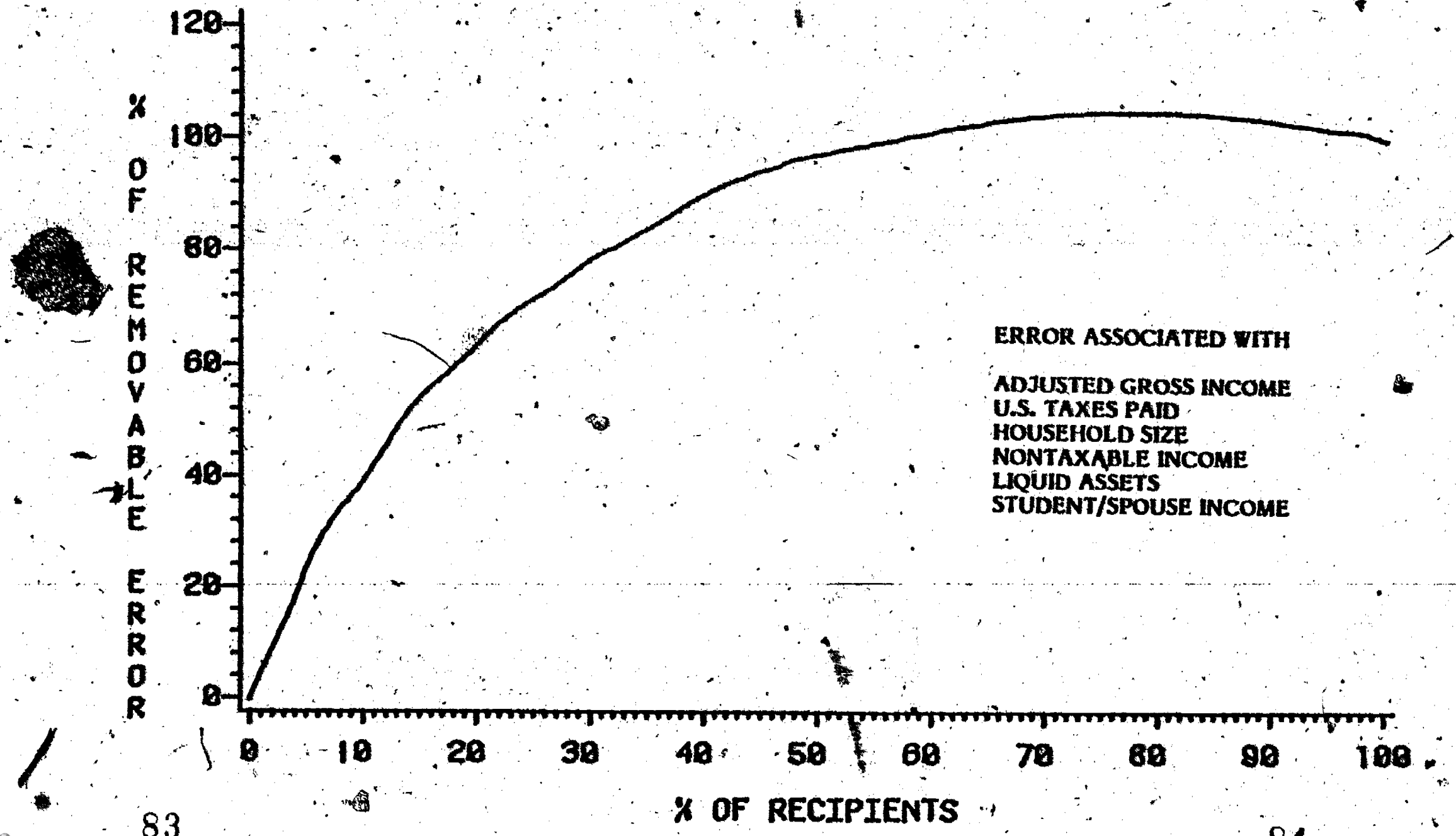
	29	41	25	43	49	23
Status	Dependent	Independent	Dependent	Dependent	Dependent	Independent
Adjusted Gross Income	Over $10,000	Over $8,000	Over $16,000	Over $20,000	$10,001 - $16,000	$8,000 and Under
Income Portion of Father/Student	Over 66%	Over 78%				Over 78%
Lived with Parents 1982	No		Yes	Yes	Yes	
Source of Tax Figures			From Tax Form	Not From Tax Form	From Tax Form	Not From Tax Form
Number of Exemptions		2 and Under	6 and Under	6 and Under	6 and Under	
Gross Income			$10,001 - $17,500		$10,001 - $17,500	
Net Household Assets						Over $2,500
Net Income						
Household Size						
Age						
Cash and Savings						
Parent's Marital Status					Married	
Home Value					$20,000 and Under	
Mean Group Error	219	208	171	150	142	123

	31	47	51	35	28	40
Status	Dependent	Independent	Dependent	Dependent	Dependent	Independent
Adjusted Gross Income	$10,000 and Under	$8,000 and Under	$10,001 - $20,000	Over $10,000	Over $10,000	Over $8,000
Income Portion of Father/Student		Over 78%			66% and Under	Over 78%
Lived with Parents 1982			Yes	Yes	No	
Source of Tax Figures		From Tax Form	Not From Tax Form	From Tax Form		
Number of Exemptions			6 and Under	6 and Under		Over 2
Gross Income				Over $17,500		
Net Household Assets	Over $20,000	$0 and Under				
Net Income		$5,000 and Under	$15,000 and Under			
Household Size	Over 3					
Age		Over 24				
Cash and Savings				$0		
Parent's Marital Status						
Home Value						
Mean Group Error	118	116	109	80	77	58

	44	26
Status	Dependent	Dependent
Adjusted Gross Income	Over $10,000	$10,000 and Under
Income Portion of Father/Student	Over 54%	
Lived with Parents 1982	Yes	
Source of Tax Figures	From Tax Form	
Number of Exemptions	6 and Under	
Gross Income	Over $17,500	
Net Household Assets		
Net Income		
Household Size		3 and Under
Age		
Cash and Savings	Over $0	
Parent's Marital Status		
Home Value		
Mean Group Error	-2	-13

REMOVABLE ERROR BY PERCENT OF RECIPIENTS

ERROR PRONE MODEL=6



GROUP NUMBER	AVERAGE NET ERROR	CUMULATIVE NET ERROR	NUMBER OF CASES	CUMULATIVE PERCENT OF CASES
39	$236	7.9	41	1.7
25	218	17.8	56	3.9
43	208	25.5	37	5.4
47	168	31.5	39	7.0
31	128	39.1	71	9.9
27	127	48.8	78	13.0
33	126	55.8	68	15.8
45	85	59.4	52	17.9
37	82	66.7	91	21.6
46	74	71.6	84	25.0
51	66	75.9	79	28.2
42	58	78.1	36	29.6
38	55	80.0	42	31.3
26	49	84.5	94	35.2
35	45	91.1	141	40.9
49	39	94.9	120	45.7
40	27	96.4	50	47.8
36	20	97.1	37	49.3
32	19	101.4	267	60.1
20	14	103.9	176	67.2
50	11	104.8	102	71.3
34	0	104.7	323	84.4
16	-7	104.4	47	86.3
28	-13	101.9	233	95.7
14	-16	101.1	69	98.5
48	-36	100.0	36	100.0

EPM 6:

AVERAGE NET STUDENT ERROR AND GROUP SIZES FOR FINAL GROUPS

EPM 6:

MEAN OVERAWARDS AND UNDERAWARDS
FOR FINAL GROUPS

GROUP NUMBER	OVERAWARD MEAN	N	UNDERAWARD MEAN	N	NO ERROR N
39	$572	17	$	0	24
25	451	30	-151	7	19
43	518	15	-99	1	21
47	310	23	-198	3	13
31	286	33	-162	4	34
27	581	17	-96	1	60
33	410	22	-182	1	45
45	236	21	-130	4	27
37	575	13		0	78
46	256	32	-163	12	40
51	238	26	-269	4	49
42	277	9	-131	3	24
38	184	15	-153	3	24
26	581	8		0	86
35	410	10	-83	3	122
49	182	40	-222	12	68
40	182	19	-247	6	25
36	216	8	-1004	1	28
32	253	28	-198	10	229
20	400	17	-370	12	147
50	177	18	-223	9	75
34	194	13	-238	11	299
16	327	6	-183	12	29
18	120	20	-183	27	184
14	284	11	-329	11	47
48	148	7	-373	6	23

	39	25	43	47	31	27
Status	Dependent	Dependent	Independent	Dependent	Dependent	Independent
Adjusted Gross Income	Over $10,000	Over $10,000	Over $8,000	Over $10,000	Over $10,000	$8,000 and Under
Income Portion of Father/Student	Over 0%	Over 84%	Over 78%	84% and Under		Over 78%
Tax Figures Source	From Tax Form	Not From Return		Not From Return	From Tax Form	Not From Return
Investment Value	$0				$0	
Taxes Paid	Over $0				Over $0	
Home Value	Over $10,000				Over $10,000	
Net Household Assets						
Net Income	$15,000 and Under				Over $15,000	Over $2,500
Household Size						
Supported by Parents '82					No	
Age				Over 21		
Exemptions			2 and Under	5 and Under		
Number in College						
Mean Error	236	218	208	168	128	127

	33	45	37	46	51	42
Status	Dependent	Dependent	Independent	Dependent	Dependent	Independent
Adjusted Gross Income	$10,000 and Under	Over $10,000	$8,000 and Under	Over $10,000	Over $10,000	Over $8,000
Income Portion of Father/Student		48% and Under	Over 78%	84% and Under		Over 78%
Tax Figures Source		From Tax Form	From Tax Form	Not From Return	From Tax Form	
Investment Value		$0			$0	
Taxes Paid		Over $0			Over $0	
Home Value		Over $10,000			$10,000 and Under	
Net Household Assets	Over $20,000		$0 and Under			
Net Income		Over $15,000	$5,000 and Under			
Household Size	Over 3					
Supported by Parents '82		Yes				
Age				21 and Under		
Exemptions				5 and Under		Over 2
Number in College					Over 1	
Mean Error	126	85	82	74	66	58

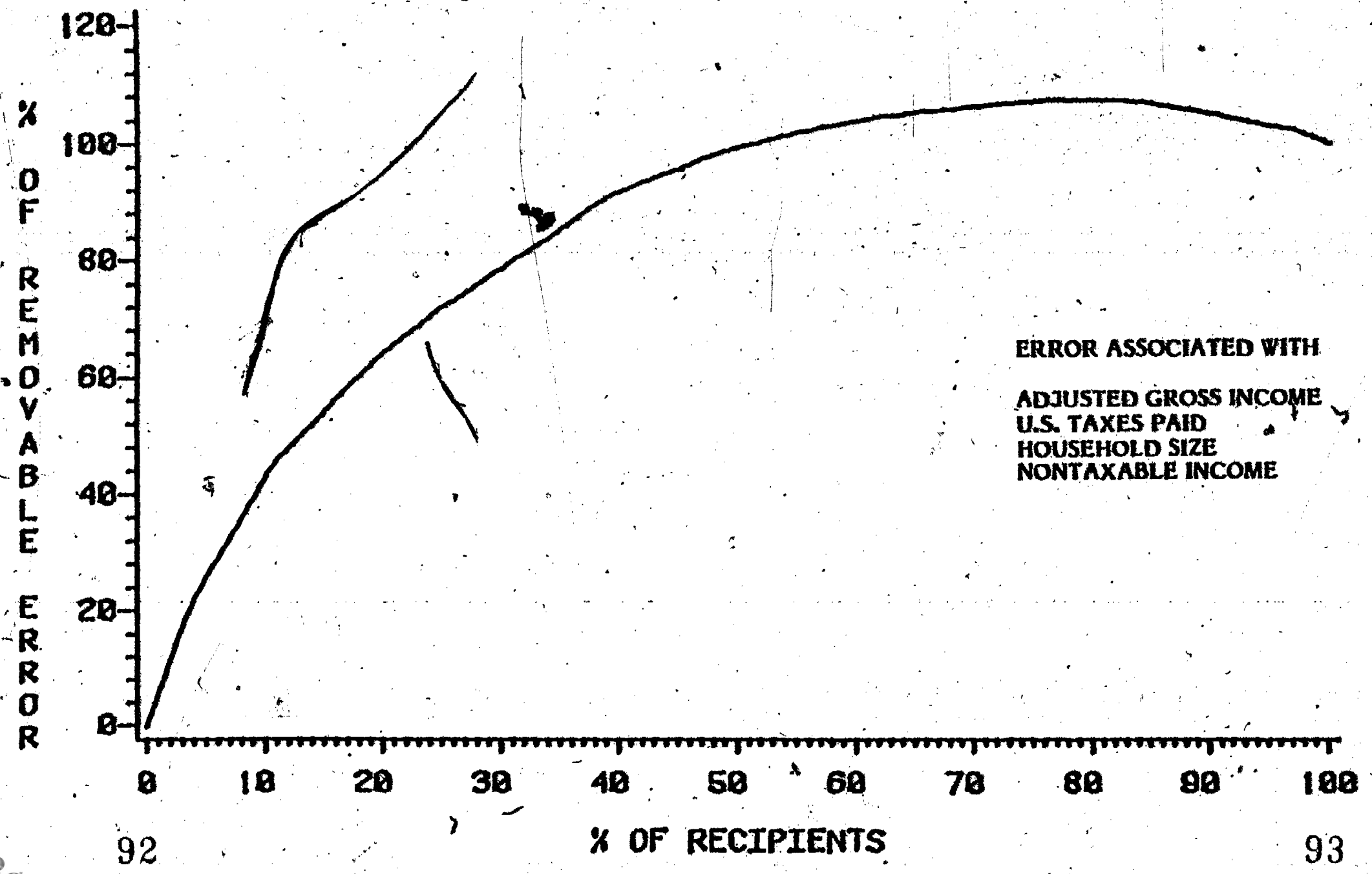
	38	26	35	49	40	36
Status	Dependent	Independent	Independent	Dependent	Dependent	Independent
Adjusted Gross Income	Over $10,000	$8,000 and Under		Over $10,000	Over $10,000	$8,000 and Under
Income Portion of Father/Student	0%	Over 78%	78% and Under	Over 60%	84% and Under	Over 78%
Tax Figures Source	From Tax Form	Not From Return		From Tax Form	Not From Return	From Tax Form
Investment Value	$0			$0		
Taxes Paid	Over $0			Over $0		
Home Value	Over $10,000			Over $10,000		
Net Household Assets						$0 and Under
Net Income	$15,000 and Under	$2,500 and Under		Over $15,000		Over $5,000
Household Size				Yes		
Supported by Parents '82						
Age			Over 30			
Exemptions					Over 5	
Number in College						
Mean Error	55	49	45	39	37	20

	32	20	50	34	16	28
Status	Dependent	Independent	Dependent	Independent	Dependent	Dependent
Adjusted Gross Income	$10,000 and Under	$8,000 and Under	Over $10,000		Over $10,000	$10,000 and Under
Income Portion of Father/Student		Over 78%		78% and Under		
Tax Figures Source		From Tax Form	From Tax Form		From Tax Form	
Investment Value			$0		$0	
Taxes Paid			Over $0		$0	
Home Value			$10,000 and Under			
Net Household Assets	$20,000 and Under	Over $0				
Net Income						
Household Size	Over 3					3 and Under
Supported by Parents '82						
Age				30 and Under		
Exemptions			1			
Number in College						
Mean Error	19	14	11	0	-7	-13

	14	48
Status	Dependent	Dependent
Adjusted Gross Income	Over $10,000	Over $10,000
Income Portion of Father/Student		49% - 60%
Tax Figures Source	From Tax Form	From Tax Form
Investment Value	Over $0	$0
Taxes Paid		Over $0
Home Value		Over $10,000
Net Household Assets		
Net Income		Over $15,000
Household Size		
Supported by Parents '82		Yes
Age		
Exemptions		
Number in College		
Mean Error	-16	-36

REMOVABLE ERROR BY PERCENT OF RECIPIENTS

ERROR PRONE MODEL-7



GROUP NUMBER	AVERAGE NET ERROR	CUMULATIVE NET ERROR *	NUMBER OF CASES	CUMULATIVE PERCENT OF CASES
27	$238	9.3	40	1.6
47	208	18.2	37	3.1
45	170	27.0	54	5.3
21	123	37.9	78	8.5
51	116	43.3	41	10.1
31	104	48.8	54	12.3
49	91	53.9	61	14.8
44	88	62.7	106	19.1
43	68	69.0	99	23.1
46	58	71.5	36	24.5
26	53	74.3	54	26.7
23	49	76.8	43	28.5
20	48	81.9	94	32.3
35	45	89.6	141	38.0
50	42	91.9	50	40.0
36	34	93.5	49	42.0
39	31	95.6	58	44.4
42	27	102.3	272	55.4
48	15	104.6	153	61.6
40	10	105.0	37	63.1
24	9	106.2	144	68.9
32	8	106.9	86	72.4
34	0	106.8	323	85.5
28	-15	103.4	231	94.8
14	-18	102.5	53	97.0
38	-29	100.0	75	100.0

EPM 7:

AVERAGE NET STUDENT ERROR AND GROUP SIZES
FOR FINAL GROUPS

EPM 7:

MEAN OVERAWARDS AND UNDERAWARDS FOR FINAL GROUPS

GROUP NUMBER	OVERAWARD MEAN	OVERAWARD N	UNDERAWARD MEAN	UNDERAWARD N	NO ERROR N
27	$549	18	$-125	2	20
47	518	15	- 99	1	21
45	346	28	-108	5	21
21	560	17		0	61
51	687	7		0	34
31	220	27	-233	1	26
49	323	22	-215	7	32
44	249	40	-108	7	59
43	323	25	-209	6	68
46	277	9	-131	3	24
26	353	15	-207	12	27
23	587	4	-170	2	37
20	570	8		0	86
35	410	16	-83	3	122
50	423	5		0	45
36	129	15	-109	2	32
39	907	2		0	56
42	164	69	-192	22	181
48	157	34	-182	16	103
[illegible]	230	6	-1004	1	30
2[illegible]	201	15	-136	12	117
32	211	6	-266	2	78
24	193	13	-238	11	299
28	157	19	-192	33	179
14	316	12	-502	10	31
38	278	6	-439	9	60

	27	47	45	21	51	31
Status	Dependent	Independent	Dependent	Independent	Independent	Dependent
Income Portion of Father/St[illegible]		Over 78%		Over 78%	Over 78%	
Household Size	Over 3		4 to 6			Over 3
Adjusted Gross Income		Over $8,000	Over $0	$8,000 and Under	$8,000 and Under	
Lived with Parents 1982	No		Yes			Yes
Tax Figures Estimated			Not From Return	Not From Return	From Tax Form	From Tax Form
Transaction Number			2 and Under			2 or Under
Net Household Assets					$0 and Under	
Net Income				Over $2,500	$5,000 and Under	
Taxes Paid						Over $0
Exemptions	4 and Under	2 and Under				3 and Under
Age					Over 24	
Student/Spouse Assets			Over $0			
Parent's Marital Status						
Mean Error	238	208	170	123	116	104

	49	44	43	46	26	23
St[illegible]	Dependent	Dependent	Dependent	Independent	Dependent	Independent
Income Portion of Father/Student			Over 84%	Over 78%		Over 78%
Household Size	3 and Under	4 - 6	Over 3		Over 3	Over 1
Adjusted Gross Income		Over $0		Over $8,000		$8,000 and Under
Lived with Parents 1982		Yes	Yes		No	
Tax Figures Estimated		Not From Return	From Tax Form			From Tax Form
Transaction Number		2 or Under	2 or Under			
Net Household Assets						Over $0
Net Income			Over $0			
Taxes Paid	Over $500		Over 3	Over 2	Over 4	
Exemptions						
Age						
Student/Spouse Assets		$0				
Parent's Marital Status	Married					
Mean Error	91	88	68	58	53	49

	20	35	50	36	39	42
Status	Independent	Independent	Independent	Dependent	Independent	Dependent
Income Portion of Father/Student	Over 78%	78% and Under	Over 78%		Over 78%	84% and Under
Household Size				Over 6	1	Over 3
Adjusted Gross Income	$8,000 and Under		$8,000 and Under	Over $0	$8,000 and Under	
Lived with Parents 1982				Yes		Yes
Tax Figures Estimated	Not From Return		From Tax Form	Not From Return	From Tax Form	From Tax Form
Transaction Number				2 and Under		2 and Under
Net Household Assets			$0 and Under		Over $0	
Net Income	$2,500 and Under		$5,000 and Under			
Taxes Paid					$0	Over $0
Exemptions						Over 3
Age		Over 30	24 and Under			
Student/Spouse Assets						
Parent's Marital Status						
Mean Error	48	45	42	34	31	27

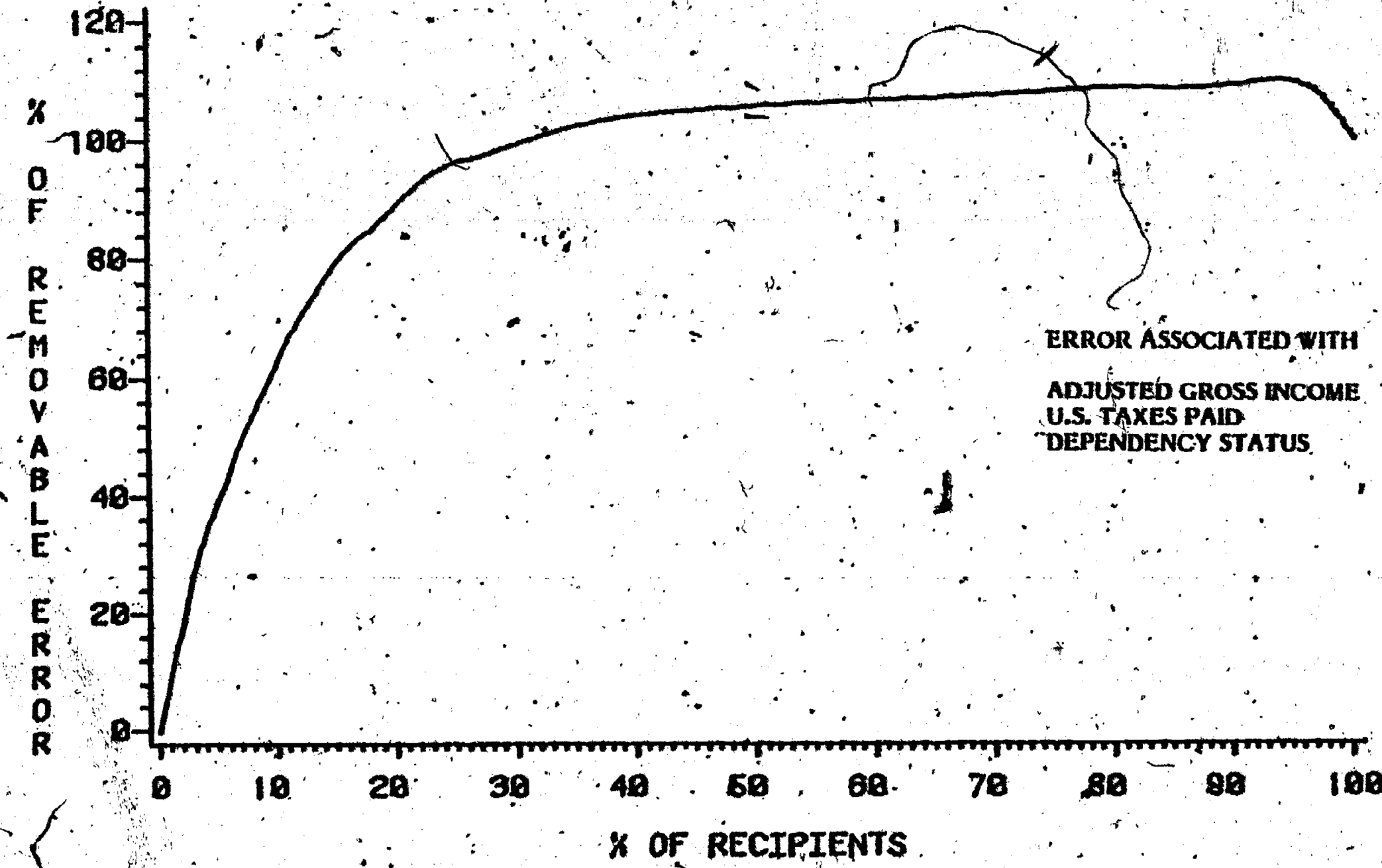
	48	40	24	32	34	28
Status	Dependent	Independent	Dependent	Dependent	Independent	Dependent
Income Portion of Father/Student		Over 78%			78% and Under	
Household Size	3 or Under		Over 3	Over 3		3 and Under
Adjusted Gross Income		$8,000 and Under		$0 and Under		
Lived with Parents 1982			Yes	Yes		
Tax Figures Estimated		From Tax Form	From Tax Form	Not From Return		
Transaction Number			2 and Under	2 and Under		
Net Household Assets		$0 and Under				
Net Income		Over $5,000				
Taxes Paid	Over $500		$0			$500 and Under
Exemptions						
Age					30 and Under	
Student/Spouse Assets						
Parent's Marital Status	All But Married					
Mean Error	15	10	9	8	0	-15

	14	38
Status	Dependent	Independent
Income Portion of Father/Student		Over 78%
Household Size	Over 3	1
Adjusted Gross Income		$8,000 and Under
Lived with Parents 1982	Yes	
Tax Figures Estimated		From Tax Form
Transaction Number	Over 2	
Net Household Assets		Over $0
Net Income		
Taxes Paid		Over $0
Exemptions		
Age		
Student/Spouse Assets		
Parent's Marital Status		
Mean Error	-2	-29

EPH

REMOVABLE ERROR BY PERCENT OF RECIPIENTS

ERROR PRONE MODEL-8



GROUP NUMBER	AVERAGE NET ERROR	CUMULATIVE NET ERROR %	NUMBER OF CASES	CUMULATIVE PERCENT OF CASES
15	$161	21.5	57	2.3
31	160	32.6	35	3.7
35	109	40.7	38	5.3
33	94	49.0	38	6.8
41	91	56.0	40	8.4
27	75	66.6	60	10.9
11	66	72.7	46	12.7
45	50	78.8	48	14.7
32	40	82.8	43	16.4
40	34	86.0	47	18.3
26	32	92.0	80	21.5
12	29	95.0	43	23.3
30	15	96.6	53	25.4
38	13	99.0	96	29.3
28	13	101.6	101	33.4
20	7	104.3	158	39.8
24	4	104.9	70	42.6
19	2	108.4	786	74.5
44	2	108.9	96	78.3
43	1	109.4	278	89.5
22	-3	108.4	178	96.7
42	-35	104.7	45	98.5
18	-66	100.0	36	100.0

EPM 8:

AVERAGE NET STUDENT ERROR AND GROUP SIZES FOR FINAL GROUPS

EPM 8:

MEAN OVERAWARDS AND UNDERAWARDS
FOR FINAL GROUPS

GROUP NUMBER	OVERAWARD MEAN	N	UNDERAWARD MEAN	N	NO ERROR N
15	$605	15	$ -	0	4
31	550	10	-32	1	24
35	359	13	-285	2	23
33	897	4		0	34
41	303	13	-91	5	22
27	498	9		0	51
11	231	13	-63	2	31
45	373	7	-101	2	39
32	577	3		0	40
40	195	10	-85	4	33
26	631	4	-12	1	75
12	208	6	-12	1	36
30	135	12	-385	2	39
38	165	18	-156	10	68
28	163	11	-158	3	87
20	174	9	-114	4	145
24	256	1		0	69
19	158	50	-187	31	705
44	183	1		0	92
43	123	5	-80	5	268
22	53	7	-177	5	166
42	82	4	-638	3	38
18	304	3	-560	6	27

	15	31	35	33	41	27
Status	Independent	Dependent	Dependent	Independent	Dependent	Independent
Income Portion of Father/Student	Over 78%			Over 78%	Over 42%	Over 78%
Tax Figures Source	Not From Return		Not From Return	Not From Return	Not From Return	From Tax Form
Net Income	Over $2,500	Over $25,000	$25,000 and Under	$2,500 and Under	$25,000 and Under	$2,501 - $5,000
Taxes Paid			$3,000 and Under		$3,000 and Under	
Household Size	1			1		
Transaction Number			Over 1		1	
Parent's Marital Status			Divorced or Married		Divorced or Married	
Net Household Assets						$0
Claimed as Exemption '82			Yes		Yes	
Itemized Deductions		$1,000 and Under				
Age				22 and Under		
Number in College					1	
Exemptions						
Mean Error	161	160	109	94	91	75

	11	45	32	40	26	12
Status	Dependent	Independent	Independent	Dependent	Independent	Independent
Income Portion of Father/Student		78% and Under	Over 78%	42% and Under	Over 78%	Over 78%
Tax Figures Source			Not From Return	Not From Return	From Tax Return	Not From Return
Net Income	$25,000 and Under		$2,500 and Under	$25,000 and Under	$2,501 - $5,000	
Taxes Paid	Over $3,000			$3,000 and Under		
Household Size			1			Over 1
Transaction Number				1		
Parent's Marital Status				Divorced or Married		
Net Household Assets					Over $0	
Claimed as Exemption '82				Yes		
Itemized Deductions						
Age		Over 30	Over 22			
Number in College				1		
Exemptions		Over 2				
Mean Error	66	50	40	34	32	29

	30	38	28	20	24	19
Status	Dependent	Dependent	Dependent	Independent	Independent	Dependent
Income Portion of Father/Student				Over 78%	Over 78%	
Tax Figures Source		Not From Return	Not From Return	From Tax Form	From Tax Form	From Tax Form
Net Income	Over $25,000	$25,000 and Under	$25,000 and Under	Over $5,000	$2,500 and Under	$25,000 and Under
Taxes Paid		$3,000 and Under	$3,000 and Under			$3,000 and Under
Household Size		1				2 and Under
Transaction Number						
Parent's Marital Status		Divorced or Married	Divorced or Married			
Net Household Assets						
Claimed as Exemption '82		Yes	No			
Itemized Deductions	Over $1,000					
Age						
Number in College		Over 1				
Exemptions						
Mean Error	15	13	13	7	4	2

	44	43	22	42	18
Status	Independent	Independent	Dependent	Independent	Dependent
Income Portion of Father/Student	78% and Under	78% and Under		78% and Under	
Tax Figures Source			Not From Return		From Tax Form
Net Income		$10,000 and Under	$25,000 and Under	Over $10,000	$25,000 and Under
Taxes Paid			$3,000 and Under		$3,000 and Under
Household Size					Over 2
Transaction Number					
Parent's Marital Status			Single, Widowed Separated, etc.		
Net Household Assets					
Claimed as Exemption '82					
Itemized Deductions					
Age	Over 30	30 and Under		30 and Under	
Number in College					
Exemptions	2 and Under				
Mean Error	2	1	-3	-35	-66

APPENDIX B

AID CODING CATEGORIES FOR PREDICTOR VARIABLES

AGEJ: Age in Years

Code	Definition	Percent In Category
0	Under 18	0.36
1	18	0.65
2	19	14.30
3	20	17.05
4	21	14.30
5	22	13.53
6	23	7.90
7	24	4.82
8	25 to 30	16.12
9	31 to 35	5.71
10	36 to 40	2.51
11	Over 40	2.75

BSFDBTJ: Business and Farm Debt

Code	Definition	Percent In Category
0	$0	95.46
1	1 to 5,000	1.58
2	5,001 to 20,000	0.69
3	20,001 to 40,000	0.97
4	Over 40,000	1.30

BSFVALJ: Business and Farm Value

Code	Definition	Percent In Category
0	$0	93.14
1	1 to 5,000	1.62
2	5,001 to 15,000	1.42
3	15,001 to 30,000	1.13
4	30,001 to 70,000	1.38
5	Over 70,000	1.30

TRNO: Number of Transactions

Code	Definition	Percent In Category
1	1	80.15
2	2	14.86
3	3	3.52
4	4	1.30
5	Over 4	0.16

CITIZEN: Citizenship

Code	Definition	Percent In Category
0	Missing or Ineligible	0.16
1	Permanent Eligible Alien	5.82
2	U.S. Citizen	94.02

DADPORTJ: Income Portion of Father/Student

Code	Definition	Percent In Category
0	0	42.85
1	.010 to .100	1.22
2	.101 to .180	1.01
3	.181 to .240	1.05
4	.241 to .300	1.09
5	.301 to .360	0.77
6	.361 to .420	1.42
7	.421 to .480	2.07
8	.481 to .540	2.15
9	.541 to .600	2.67
10	.601 to .660	2.47
11	.661 to .720	2.11
12	.721 to .780	2.27
13	.781 to .840	4.17
14	.841 to .900	10.77
15	.901 to .990	21.26
16	1	0.65

FILED: Whether or Not Taxes were Filed

Code	Definition	Percent In Category
0	Did Not File	18.08
1	Missing	0.85
2	Estimated	1.06
3	Filed	80.01

GROSINCJ: The Sum of AGI, Social Security and Other Non-Taxable Income

Code	Definition	Percent In Category
0	Less than $0	3.89
1	1 to 2,500	10.90
2	2,501 to 5,000	16.57
3	5,001 to 7,500	13.69

GROSINCJ: The Sum of AGI, Social Security and Other Non-Taxable Income (continued)

Code	Definition	Percent In Category
4	7,501 to 10,000	9.48
5	10,001 to 12,500	10.49
6	12,501 to 15,000	9.73
7	15,001 to 17,500	6.36
8	17,501 to 20,000	6.03
9	20,001 to 25,000	8.87
10	25,001 to 30,000	4.29
11	30,001 to 40,000	1.94
12	Above 40,000	0.12

HOMDBTJ: Home Debt

Code	Definition	Percent In Category
0	$0	69.38
1	1 to 5,000	4.94
2	5,101 to 10,000	7.01
3	10,001 to 15,000	6.03
4	15,001 to 20,000	3.77
5	20,001 to 25,000	2.88
6	25,001 to 30,000	1.90
7	30,001 to 35,000	1.22
8	35,001 to 40,000	0.89
9	40,001 to 45,000	0.97
10	45,001 to 50,000	0.32
11	Over 50,000	0.69

HOMEEQJ: Home Debt Subtracted from Home Value

Code	Definition	Percent In Category
0	Less Than $0	0.12
1	0	60.47
2	1 to 5,000	6.24
3	5,001 to 10,000	8.30
4	10,001 to 15,000	5.87
5	15,001 to 20,000	5.51
6	20,001 to 25,000	4.90
7	25,001 to 30,000	3.85
8	30,001 to 35,000	2.03
9	35,001 to 40,000	1.42
10	40,001 to 45,000	0.61
11	Over 45,000	0.69

HOMVALJ: Home Value

Code	Definition	Percent In Category
0	$0	59.90
1	1 to 5,000	2.11
2	5,001 to 10,000	3.73
3	10,001 to 15,000	3.20
4	15,001 to 20,000	5.63
5	20,001 to 25,000	4.01
6	25,001 to 30,000	5.63
7	30,001 to 35,000	4.05
8	35,001 to 40,000	3.69
9	40,001 to 45,000	2.43
10	45,001 to 50,000	2.23
11	50,001 to 60,000	1.58
12	60,001 to 70,000	1.13
13	Above 70,000	0.69

INVDBTJ: Investment Debt

Code	Definition	Percent In Category
0	$0	92.73
1	1 to 5,000	0.77
2	5,001 to 15,000	0.77
3	Over 15,000	0.73

INVSTEQJ: Investment Debt Subtracted from Investment Value

Code	Definition	Percent In Category
0	Less than $0	-
1	0	94.29
2	1 to 5,000	3.08
3	5,001 to 15,000	1.98
4	Over 15,000	0.65

INVVALJ: Investment Value

Code	Definition	Percent In Category
0	$0	94.13
1	1 to 5,000	2.35
2	5,001 to 10,000	1.34
3	10,001 to 30,000	1.62
4	Over 30,000	0.57

HHSZ: Household Size

Code	Definition	Percent In Category
1	1	20.05
2	2	14.18
3	3	17.86
4	4	16.44
5	5	13.89
6	6	7.74
7	7	4.41
8	8	3.00
9	9	1.26
10	Over 9	1.17

NPHE: Number Enrolled in Postsecondary Education

Code	Definition	Percent In Category
1	1	69.30
2	2	22.68
3	3	7.09
4	Over 3	0.93

NETINCJ: Taxes Subtracted from Gross Income

Code	Definition	Percent In Category
0	Less Than $0	3.89
1	1 to 2,500	11.06
2	2,501 to 5,000	17.09
3	5,001 to 7,500	13.57
4	7,501 to 10,000	10.49
5	10,001 to 12,500	10.94
6	12,501 to 15,000	7.53
7	15,001 to 17,500	7.70
8	17,501 to 20,000	6.28
9	20,001 to 25,000	7.82
10	25,001 to 30,000	2.88
11	30,001 to 40,000	0.65
12	Over 40,000	0.12

NHAJ: Sum of Savings, Home Equity, Investment Equity

Code	Definition	Percent In Category
0	Less Than $0	0.12
1	0	32.77
2	1 to 5,000	30.82
3	5,001 to 10,000	7.94
4	10,001 to 15,000	5.79
5	15,001 to 20,000	5.43
6	20,001 to 25,000	5.27
7	25,001 to 30,000	4.13
8	30,001 to 35,000	2.67
9	35,001 to 40,000	2.35
10	40,001 to 45,000	1.13
11	45,001 to 50,000	0.69
12	50,001 to 60,000	0.73
13	Over 60,000	0.16

NVBFJ: Business and Farm Debt Subtracted from Business and Farm Value

Code	Definition	Percent In Category
0	Less Than $0	0.08
1	0	93.40
2	1 to 5,000	1.98
3	5,001 to 10,000	1.05
4	10,001 to 20,000	1.17
5	20,001 to 40,000	1.46
6	Over 40,000	0.85

PARMAR: Parents' Marital Status

Code	Definition	Percent In Category
0	Independent	40.79
1	Single	1.86
2	Married	33.94
3	Divorced	11.18
4	Widowed	6.64
5	Separated	5.43
6	Other	0.16

STATUS: Independent, Dependent Status of Student

Code	Definition	Percent In Category
0	Dependent	58.97
1	Independent	41.03

STDASTJ: Student Assets

Code	Definition	Percent In Category
0	$0	82.06
1	1 to 500	12.52
2	501 to 1,000	3.08
3	1,001 to 1,500	1.05
4	1,501 to 15,000	0.93
5	Over 15,000	0.36

STUDMAR: Marital Status of the Student

Code	Definition	Percent In Category
0	Student Presumed Single	0.89
1	Unmarried	84.85
2	Married	9.32
3	Missing	4.94

TAXFIG: Source of Tax Figures

Code	Definition	Percent In Category
0	Estimated	40.58
1	Filed	59.42

TAXPAIDJ: Taxes Paid

Code	Definition	Percent In Category
0	$0	46.05
1	1 to 500	18.15
2	501 to 1,000	11.10
3	1,001 to 1,500	7.45
4	1,501 to 2,000	5.91
5	2,001 to 2,500	4.37
6	2,501 to 3,000	3.08

TAXPAIDJ: Taxes Paid (continued)

Code	Definition	Percent In Category
7	3,001 to 3,500	2.07
8	3,501 to 4,000	0.89
9	4,001 to 4,500	0.36
10	4,501 to 5,000	0.24
11	5,001 to 7,000	0.28
12	Over 7,000	0.04

TUITIONJ: Unreimbursed Elementary and Secondary School Tuition

Code	Definition	Percent In Category
0	$0	91.66
1	1 to 200	2.84
2	201 to 400	1.26
3	401 to 600	1.01
4	601 to 800	0.81
5	801 to 1,000	0.89
6	1,001 to 1,600	0.57
7	Over 1,600	0.97

UMEDOLRJ: Unusual Medical Expenses in Dollars

Code	Definition	Percent In Category
0	$0	98.14
1	1 to 200	1.30
2	201 to 400	0.49
3	401 to 1,600	0.68
4	Over 1,600	

UMEPERCJ: Unusual Medical Expenses - Percent of Net Income

Code	Definition	Percent In Category
0	Net Income Under $0	0.69
1	No Medical Expenses	98.14
2	1 to 10%	1.01
3	11 to 30%	0.12
4	Over 30%	0.04

XMPT: Number of Exemptions

Code	Definition	Percent In Category
0	0	22.80
1	1	15.96
2	2	11.30
3	3	12.84
4	4	11.66
5	5	10.98
6	6	6.64
7	7	3.44
8	8	2.35
9	9	1.13
10	Over 9	0.89

AGI: Adjusted Gross Income

Code	Definition	Percent In Category
0	Under $0	1[illegible].94
1	1 to 2,000	8.59
2	2,001 to 4,000	9.84
3	4,001 to 6,000	8.75
4	6,001 to 8,000	7.53
5	8,001 to 10,000	6.64
6	10,001 to 12,000	7.70
7	12,001 to 14,000	5.22
8	14,001 to 16,000	4.86
9	16,001 to 18,000	4.37
10	18,001 to 20,000	4.62
11	20,001 to 25,000	7.90
12	25,001 to 30,000	3.97
13	30,001 to 40,000	1.94
14	Over 40,000	0.12

SOCSEC: Social Security

Code	Definition	Percent In Category
0	$0	86.84
1	1 to 1,000	1.13
2	1,001 to 2,000	1.26
3	2,001 to 3,000	1.46
4	3,001 to 4,000	2.75
5	4,001 to 5,000	1.46
6	5,001 to 6,000	1.38
7	6,001 to 7,000	0.93
8	7,001 to 8,000	0.61

SOCSEC: Social Security (continued)

Code	Definition	Percent In Category
9	8,001 to 10,000	1.34
10	Over 10,000	0.85

NONTAX: Non-Taxable Income Excluding Social Security

Code	Definition	Percent In Category
0	$0	75.62
1	1 to 2,000	10.77
2	2,001 to 4,000	8.14
3	4,001 to 6,000	3.60
4	6,001 to 8,000	1.17
5	Over 8,000	0.69

STUDINC: Student and Spouse Income

Code	Definition	Percent In Category
0	$0	72.54
1	1 to 1,000	8.55
2	1,001 to 2,000	9.23
3	2,001 to 3,000	5.18
4	3,001 to 4,000	2.71
5	4,001 to 9,000	1.22
6	9,001 to 15,000	0.20
7	Over 15,000	0.36

ITEM: Itemized Deductions

Code	Definition	Percent In Category
0	$0	83.07
1	1 to 1,000	1.17
2	1,001 to 2,000	0.93
3	2,001 to 3,000	1.86
4	3,001 to 4,000	3.60
5	4,001 to 5,000	4.17
6	5,001 to 6,000	1.86
7	6,001 to 7,000	1.34
8	7,001 to 8,000	0.61
9	8,001 to 9,000	0.53
10	Over 9,000	0.85

LDV79: Lived with Parents in 1981

Code	Definition	Percent In Category
0	No	43.05
1	Yes	56.96

CLMD79: Claimed by Parents for Tax Purposes in 1981

Code	Definition	Percent In Category
0	No	52.77
1	Yes	47.23

ASSTD79: Assisted Financially by Parents in 1981

Code	Definition	Percent In Category
0	No	61.32
1	Yes	38.68

LVD80: Lived with Parents in 1982

Code	Definition	Percent In Category
0	No	46.50
1	Yes	53.50

CLMD80: Claimed by Parents for Tax Purposes in 1982

Code	Definition	Percent In Category
0	No	55.45
1	Yes	44.55

ASSTD80: Assisted Financially by Parents in 1982

Code	Definition	Percent In Category
0	No	64.28
1	Yes	35.72

VABENJ: Veterans Benefits

Code	Definition	Percent In Category
0	0	96.84
1	1 to 1,000	0.20
2	1,001 to 3,000	1.46
3	3,001 to 4,000	0.69
4	Over 4,000	0.81

STSOCJ: Student's Projected Social Security for 1982

Code	Definition	Percent In Category
0	0	91.90
1	1 to 1,000	1.78
2	1,001 to 1,500	1.42
3	1,501 to 2,000	1.17
4	2,001 to 2,500	1.09
5	2,501 to 3,000	1.09
6	3,001 to 3,500	0.61
7	Over 3,500	0.93

SAIOJ: SAI Using Inflated Computed Applicant Record (CAR) Figures and 1982-83 Computation Formula

Code	Definition	Percent In Category
0	0	43.30
1	1 to 100	5.83
2	101 to 200	4.37
3	201 to 300	4.21
4	301 to 400	3.77
5	401 to 500	3.32
6	501 to 600	3.97
7	601 to 700	4.01
8	701 to 800	3.40
9	801 to 900	3.12
10	901 to 1,000	3.28
11	1,001 to 1,100	3.40
12	1,101 to 1,200	2.79
13	1,201 to 1,300	2.71
14	1,301 to 1,400	2.51
15	1,401 to 1,500	2.67
16	1,501 to 1,600	2.11
17	Over 1,600	1.22

PRJINCJ: Projected Student Income for 1982

Code	Definition	Percent In Category
0	0	44.15
1	1 to 1,000	20.13
2	1,001 to 2,000	10.90
3	2,001 to 3,000	8.51
4	3,001 to 4,000	4.25
5	4,001 to 5,000	2.47
6	5,001 to 6,000	2.03
7	6,001 to 7,000	1.30
8	7,001 to 8,000	1.66
9	8,001 to 9,000	1.01
10	9,001 to 11,000	1.54
11	11,001 to 13,000	1.09
12	Over 13,000	0.97

SAVINGJ: Cash and Savings

Code	Definition	Percent In Category
0	0	45.12
1	1 to 1,000	40.14
2	1,001 to 2,000	5.22
3	2,001 to 3,000	2.63
4	3,001 to 4,000	1.42
5	4,001 to 5,000	1.05
6	5,001 to 6,000	1.05
7	6,001 to 9,000	1.30
8	9,001 to 15,000	0.93
9	Over 15,000	1.13