DOCUMENT RESUME

ED 254 172 HE 018 138

TITLE Analysis of Error Associated with the Application and Allocation Aspects of the Campus-Based Programs: Results from Initial Data Collection.

INSTITUTION Advanced Technology, Inc., Reston, VA.
SPONS AGENCY Office of Student Financial Assistance (ED), Washington, DC.

PUB DATE May 84
NOTE 39p.; This report supercedes "Report on the Analyses of Error in the FISAP After the Initial Data Collection Phase, April 25, 1984." For related documents, see HE 018 112-135 and HE 018 137-140.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Accountability; Computation; *Data Collection; Enrollment Rate; *Error Patterns; *Federal Aid; Financial Aid Applicants; Grants; Higher Education; *Quality Control; *Records (Forms); Resource Allocation; *Student Financial Aid; Student Loan Programs; Work Study Programs
IDENTIFIERS *Campus Based Financial Aid; College Work Study Program; Department of Education; National Direct Student Loan Program; Supplemental Educational Opportunity Grants

ABSTRACT

Error analysis for the institutional process of applying for federal aid is considered as part of the Quality Control Study of the U.S. Office of Student Financial Aid. Attention is focused on the results of the data collection activities that occurred using the Department of Education's records for the Fiscal-Operations Report and Application to Participate (FISAP), with attention to in-house data collection. After a description of the application and allocation processes, consideration is given to the types of error measures, including the likelihood of occurrence; severity, or size, of error; need consequence; and allocation consequence. The following types of comparisons that were used for each data element are also discussed: conceptually exact, cross-year, internal consistency, and cross-school. For each data element, a chart shows the error measures and the type of data sources used in the comparisons. Quantitative and qualitative methods of causal analysis and data collection procedures and problems are also addressed, along with the guidelines for performing the comparisons. Finally, the results of the error measurement analysis are summarized. Included is the form used to record information at the institution, as well as sample FISAP forms for specific institutions. (SW)

ANALYSIS OF ERROR ASSOCIATED WITH THE APPLICATION AND ALLOCATION ASPECTS OF THE CAMPUS-BASED PROGRAMS:

RESULTS FROM INITIAL DATA COLLECTION

MAY 1984

ANALYSIS OF ERROR ASSOCIATED WITH THE APPLICATION AND ALLOCATION ASPECTS OF THE CAMPUS-BASED PROGRAMS: RESULTS FROM INITIAL DATA COLLECTION

1.0 INTRODUCTION

Advanced Technology is conducting a Title IV Quality Control Study for the Office of Student Financial Assistance (OSFA). One task in the study is an analysis of error in the institutional process of applying for aid (the FISAP process). This analysis will use data collected in two phases: one at ED, and one at the institutions. This is an interim report presenting the results of the data collection activities that took place using records available at the Department of Education (ED) prior to the institutional data collection. This paper uses the error measures and analysis procedures described in the previous report, "Analyses of Error Associated with the Application and Allocation Aspects of the Campus-Based Programs," as a base from which to report on the analysis of the data collected.

The report has six sections. Section 2 is a description of the FISAP and allocation process. Section 3 describes the types of error measures used and the types of comparisons used for each data element. Section 4 describes the data collection procedures that were used and the problems that were associated with those procedures. Section 5 recaps the kinds of analyses that were performed and the results and limitations of the analyses. The last section presents a summation of the results.

2.0 FISAP PROCESS AND POTENTIAL ERRORS

The purposes of the FISAP are found in the two components of the form. The Fiscal Operations Report presents an accounting of funds in the three Campus-Based programs. The institutions report how much money was received, how much was spent, how it was spent, and any changes in the status of their accounts. The Application part, the focus of this analysis, contains information needed to compute the institution's allocation of funds. The information includes:

- The number of eligible financial aid applicants by income level, undergraduate/graduate status, and dependency status
- The enrollment by undergraduate/graduate status of traditional institutions and the number of continuing and new students by month for those institutions with non-traditional calendars
- Amount of tuition and fee revenue by undergraduate/graduate status
- Amount of Pell funds expended.
- The amount of state and institutionally administered grant funds expended.

2.1 OVERVIEW OF THE APPLICATION PROCESS AND POTENTIAL ERRORS

The process is, on the face, relatively simple. The data submitted by institutions is edited at least twice by ED and is checked by the institution. However, in reality the process is complex and has several points in it where errors can and do occur. These critical points in the process include:

- The compilation of information at the institution
- The actual filling out of the Application portion of the FISAP
- The data entry of the information at the contractor site
- The edit checks conducted on the data by ED
- The sending to the institution a report of the errors it has found
- The correcting/changing of information at the institution
- The editing and updating of the new data submitted by institutions
- The computation of the tentative and final allocations.

In addition, at various points in the overall process, the FISAP can be:

- Lost
- Misplaced
- Not completely filled out by the institution.

While these latter points are certainly causes of concern for ensuring the smooth operation of the process, this report is restricted to errors in the data elements on the application and the processing, and not on the receipt control and filing procedures at the institution or ED.

There are many different types of errors that can be made. Some are easily detected by computer edits such as arithmetic mistakes; others are not so easily found because there is only one source for that particular piece of information. The types of errors that can occur include:

- Compiling the wrong data for entry onto the FISAP
- Arithmetic mistakes
- Incorrect transcription of information onto the FISAP
- Incorrect data entry
- Incomplete checking of data elements at the institution
- Incomplete edits
- Incorrect allocation algorithm.

The application requires data from many sources other than the financial aid office. Compiling the correct information [redacted] requires coordination between two or more offices. This could lead to misunderstanding of definitions, lack of adequate records, and so on. The figure the financial aid office receives may not be the appropriate one, and the financial aid administrator would probably not check its validity.

The application is complex to fill out even by the most knowledgeable of aid administrators. Because the definitions on the application change periodically, and because there are internal instructions, there is a good chance that transcription errors and arithmetic errors could occur. Most of these should be detected either by the computer edits at ED or through institutional review of the error report it receives from ED.

Every data preparation contractor is required to meet a certain level of accuracy in its data entry activity. Even meeting the contractual obligation still means that errors in data entry can occur. These errors should be caught either by the computer edits or by the institution.

Perhaps the most difficult types of errors to find are those that are embedded within the software of a computer. In the FISAP process, the data that is submitted to ED is subject to about 150 computer edits that test mainly for range of values and internal consistency. Also, the institutional allocations are automatically generated. Both of the processes are complex to perform; the software systems must be thoroughly tested.

The next section goes into more detail about the errors that can occur and the description of the error measures that are used in the analysis. Not all types of errors are addressed in this interim report where the only information available at ED was reviewed and analyzed. The results of the on-site data collection, to be reported at a later date, will expand the analysis and the issues that can be addressed.

2.2 ALLOCATION ALGORITHM

The form of each of the institutional allocation funds is essentially an algorithm of subtracting institutional resources from institutional costs. There are separate formulas for SEOG and for CW-S and NDSL because the former is a grant program and the latter are self-help programs. The two formulas are:

SEOG need = .75 x ((avg. tuition and fees for undergraduates) + (avg. living cost)) x (the number of eligible aid applicants) - aggregate family contribution - Pell expenditures - state expenditures - .25 x institutional expenditures

CW-S and NDSL need = ((avg. tuition and fees for undergraduates, graduates) + (avg. living cost of undergraduates, graduates) x (number of undergraduate, graduate aid applicants) - (aggregate family contribution)

The components of the formulas are all self-explanatory except for the aggregate family contribution. For each income category of eligible aid applicants, a

family contribution amount is assigned. This amount differs across level of enrollment and dependency status. To calculate the aggregate family contribution for an institution, the number of eligible aid applicants at each income level is multiplied by the average family contribution and those products are then summed first across income groups and then across level of enrollment and dependency status.

The result of these calculations is a need figure for SEOG and CW-S/NDSL. These amounts are then added across institutions within the states to obtain a total state need for either SEOG or CW-S/NDSL. The institution's need figure is divided by the total for the state resulting in the institution's percentage of total state need. The allocation is based on this porportion as well as the proportion of an institution's need to the total need for all schools nationwide.

3.0 ERROR MEASURES, COMPARISONS, AND ANALYSIS PROCEDURES

In the previous section we describe the data elements and algorithms to allocate Campus-Based funds to individual campuses. This section presents three description of the types of analysis, measures, and comparisons to be performed for the application and allocation process.

3.1 TYPES OF ERROR MEASURES

For any particular data element there are four possible error measures. It will not be possible to utilize all four error measures for each data element. These four error measures are:

- Likelihood of occurrence
- Severity, or size, or error
- Need consequence
- Allocation consequence.

These four measures are briefly defined in the following paragraphs, and are used in the summary figure to indicate the types of measures to be estimated for each data element.

Likelihood of occurrence is simply the proportion, or percentage, of cases (schools) for which a difference between the "application" value and the "comparison" value occurs and is greater than some tolerance. Estimation of this type of error measure is feasible for all data elements and comparison values.

Severity or size of difference in error provides an estimate of the seriousness of differences between application values and comparison values. These measures would be reported as the average error or average difference between the two values. As with the occurrence likelihood measure, severity can be estimated for all data elements and comparison values.

Need consequence measures estimated effects of data element errors on the measures of absolute aggregate need calculated for each campus. This type of measure will only be used for data elements for which the comparison values represent exact values. For example, it is possible to estimate the need consequence of differences between reported enrollment and HEGIS enrollment; however, the comparison between the institutionally reported income distributions and average income distributions would not support estimation of need consequence.

Allocation consequence is the next step after need consequence and would only include data elements for which exact comparisons are supported. Allocation consequences may not follow the same pattern as need consequences because of various hold-harmless and maintenance-of-effort considerations, and because of the pattern of error across other schools.

3.2 TYPES OF COMPARISONS

There are four types of comparison values which will be used in this analysis:

- Conceptually exact
- Cross-Year
- Internal Consistency
- Cross-School.

Brief descriptions of these types of comparisons are presented in the following paragraphs. In Section 3.3 these comparison types and the measure types are brought together in a figure which summarizes the plans for analysis of errors associated with the application and allocation process.

Conceptually exact comparisons involve data elements where an alternative data source contains conceptually-comparable values for the FISAP data element. Examples would be undergraduate enrollment and Pell expenditures. For data elements which have conceputally exact values, it is possible to utilize all four types of error measures.

Cross-Year comparisons involve values from the October 1981 FISAP. These comparisons are not exact but would be expected to identify potential error when cross-year changes exceed reasonable tolerances. We would assess the likelihood and severity of cross-year differences but would not estimate need consequences or allocation consequences.

Internal Consistency comparisons would involve relationships between data elements in the application section of the FISAP. Typically, these comparisons would involve reported totals and calculated sums of components. Likelihood and severity would be reported for all internal consistency comparisons. Whether need and allocation consequences are estimable needs to be decided on a case-by-case basis.

Cross-school comparisons focus of whether or not the value reported by a school is within a reasonable range of what was reported by similar schools. It is an inexact comparison but the distributions of these differences may highlight data elements needing corrective actions. Only severity of occurrence error measures would be appropriate given the inexact nature of these types of comparisons.

3.3 DATA ELEMENTS, ERROR MEASURES, AND COMPARISONS

Figure 1 summarizes the planned error measurement analysis. It contains a row for each application data item in the allocation process. The entries in each row indicate either the type of error measure or the type of comparison to be performed for each data element.

DATA ELEMENT	ERROR MEASURES				NATURE OF COMPARISONS					
	Occurrence	Severity	Need	Allocation	X-Year FISAP	HEGIS	Pell	Internal Consist.	Occup. Survey	X-School FISAP
Maintenance of Effort	X	X			X					
Undergraduate Tuition & Fees	X	X				X				
Graduate Tuition & Fees	X	X				X				
Pell Expenditures	X	X	X	X			X			
State Expenditures	X	X			X					
Institutional Expenditures	X	X			X					
Undergraduate Dependent Applicants	X	X			X			X		X
Graduate Dependent Applicants	X	X			X			X		X
Undergraduate Independent Applicants	X	X			X			X		X
Graduate Independent Applicants	X	X			X			X		X
Undergraduate Enrollment	X	X	X	X		X				
Graduate Enrollment	X	X	X	X		X				
Continuing Enrollment	X	X			X				X	
New Enrollment	X	X			X				X	

FIGURE 1

SUMMARY OF FISAP ANALYSES

The figure highlights the diversity of analyses, the data sources used, and their relationship. In some cases, due to limitations of directly comparable data, only occurrence and severity measures of error can be constructed. In others, the analysis can pinpoint the eventual effect on an institution's allocation caused by an error in a data item.

3.4 CAUSAL ANALYSIS

There are two types of causal analysis planned in order to assess potential causes and to identify possible corrective actions. The first type is the traditional, quantitative assessment, while the second involves a more qualitative approach.

3.4.1 Quantitative Methods of Causal Analysis

This type of analysis primarily consists of bivariate cross-tabulations of data-element error-measures by institutional characteristics. The nature of the reasearch questions will be whether or not the likelihood of occurrence, severity or size of differences, and need or allocation consequences differ across the characteristics of the schools. Existence of significant differences would suggest that the associated institutional characteristics (or correlates of that characteristic) caused the error.

Institutional characteristics to be investigated as causes or correlates of error would include:

- Type of institution
- Institutional control
- Institutional office responsible for FISAP preparation
- Method of preparation used
- Size of institution
- Use of QC procedures
- Characteristics of financial aid office
- Size of Campus-Based allocation
- Federal program participation.

3.4.2 Qualitative Methods of Causal Analysis

Our proposed approach involves soliciting information directly from the institution as to the causes of any errors we uncover using the comparisons and measures discussed earlier. In order to optimally utilize scarce resources, attention will be focused on the "exact" comparisons. For any difference discovered during our previsit data collection, we will ask the FAA to provide an explanation, rationalization, etc. Thus at any school this set of questions will be tailored to the list of errors discovered for that school.

These open-ended responses will be examined by project analysts in order to distill possible causes of the errors and to suggest corrective actions.

4.0 DATA COLLECTION PROCEDURES

The data collection activities for analyzing the FISAP are planned in two steps. The first step utilizes the information that is available at the Department in the form of program data or other files. The second step is to ask the institutions directly for the information they supplied on the FISAP and how they arrived at it. Using this approach, the first step of data collection could be done before the institutional site visits. Any apparent errors discovered in the first step of data collection could then be explained by the visit to the institution. This interim report is confined to the results of the first data collection.

4.1 DATA COLLECTION PLAN

The data collection activities were restricted to those sources that were readily available at ED. Figure 1 from the previous section shows the data elements under analysis and the types of analysis to be used. Data from the following sources were collected:

- Pell Universe file
- Higher Education General Information Survey (HEGIS) Fall Enrollment survey
- HEGIS Financial Statistics survey

- October, 1981 FISAP
- Postsecondary Schools with Occupational Programs survey
- Audits and program reviews
- Hard copy of the October, 1982 FISAP.

The data collection plan was to gather the values of the relevant data items from each data source. This was done, depending upon the kind of data, in the following ways:

- Abstracting from general institutional files such as audits and program reviews
- Extracting data from program files such as the Pell Universe and the two FISAPs
- Extracting data from institutional surveys such as the HEGIS and the Postsecondary Schools with Occupational Programs surveys.

Once all the data had been gathered, the values were entered onto summary data sheets, as shown in Figure 2.

The summary data sheets contain all the information necessary to perform the initial stage of analysis. The data items of interest are listed down the left-hand side and the different sources of information listed across the top. Each occurrence of a data item for which there is a comparable value was then entered in the appropriate box. This was done for each of the 275 institutions in the sample.

4.2 IN-HOUSE DATA COLLECTION

The actual data collection procedures are complex because of the multiple data sources being utilized and because of the varied methods of analysis. Methods of data collection varied according to the form in which the data are found. For instance, the Pell Universe is on a computerized file and updated periodically so that form was the most efficient to use. The forms of the data we used are:

- Pell Universe file - computerized data file with one record for each institution

- HEGIS Fall Enrollment survey - hard copy of the HEGIS form for each institution
- HEGIS Financial Statistics survey - hard copy of the HEGIS form
- October 1981 FISAP - computerized data file with one record for each institution
- Postsecondary Schools with Occupational Programs survey - book containing data on the schools in the survey
- Audits - partially in folders and partially on microfiche with one folder or piece of microfiche for all the audits for that institution
- Program reviews - microfiche with all program reviews on the microfiche
- October 1982 FISAP - both the computerized data file and the hard copies from the ED files.

The data collection went according to the plan in most respects. In most cases, the paper and microfiche files were in fairly good order and easily accessible. The computerized files were very easy to work with. The major problem with the paper files are their bulkiness and the fact they contain far more than one year of data. Moreover, the paper files are difficult to keep track of in spite of attempts to maintain a record of when files are taken out. The purposes of the files differ across the various data files. As a result, the level of knowledge of the files and the accessibility varies.

The two major problems with this data collection were the gathering of the most recent data, and correctly identifying the school we wanted. The first problem is endemic in any program because information is constantly updated. The second problem has been a perennial one in OSFA.

All of the data files we examined with the possible exception of the Pell Universe begins with a hard copy form that is filled out, key entered, edited, and then updated at least once. Since our data collection examined both hard copy and computerized files, it would be difficult to determine where any errors exist. For example, the HEGIS Fall Enrollment survey is keyed, edited, and then changed based on additional data collection. A revision made by the school may or may not be found in the hard copy file; it may be directly keyed in. Unfortunately, using the computerized file may not be the most accurate because revisions may be made to the

hard copy, but not keyed. All of the data files used in this data collection are subject to these potential errors.

The problem of identifying the correct school is a complex one in this data collection. The problems of OSFA are compounded by the fact that our sample of institutions is not as well defined as we needed. The identification problem has two components: no common identifier across data files; and inconsistent reporting units across data files. The two types of problems overlap because while there may be a common identifier between two data files, there is no guarantee that multi-campus schools report in the same grouping across files. The Pell and Campus-Based programs have different identification numbers for the schools that participate in the respective programs. Our data collection used the entity number (EIN) as a common identifier. For the other data files we had to rely on the name of the school as the primary identifier. This caused problems when schools changed names or when different names are used for the different programs. This may have contributed to some of the problems discussed in the next section. In much the same way, schools report differently to different offices of ED. The most common occurrence of this problem is with branch campuses and consortia of schools. There are no rules for defining a reporting unit so any attempt to identify particular school's information from different data sources may be impossible. This, too, contributed to the missing data problems.

4.3 PROBLEMS

A number of problems were encountered in this phase of the data collection. The two major types of problems include:

- Missing data elements, either the documents could not be found or the data element was missing from the file
- The data values found in the different sources did not compare as closely to the definition used in the application and allocation process as was anticipated.

Missing data occurred in two forms. In the first form, data elements were either not aggregated in the same way in the alternative source as on the FISAP, or the particular data value was missing. This type of missing data occurred with the Pell Universe and HEGIS files because schools use different reporting units for different

DATA ELEMENT	FISAP	FISAP Form	HEGIS	Occup Survey	Pell	Audit	Program Review	X-Year FISAP	X-School FISAP
Maint. of Effort (II.D.9)									
UG Tuition & Fees (II.E.15a)									
Grad Tuition & Fees (II.E.15b)									
Pell $ (II.E.16)									
State $ (II.E.17)									
Institution $ (II.E.18)									
UG DEP. Applicants (II.F.35a)									
Grad Dep. Applicants (II.F.35b)									
UG Ind. Applicants (II.F.35c)									
Grad Ind. Applicants (II.F.35d)									
UG Enrollment (II.G.36a)									
Grad Enrollment (II.G.36b)									
Continuing Enrollment (II.G.51a)									
New Enrollment (II.G.51b)									

Institution: _____

FIGURE 2

SUMMARY DATA SHEET

programs. The second type of missing data was where there were no documents or trace of data. This happened primarily when we were examining the hard copies of documents for the data values. This type of missing data was found in the hard copy FISAP, HEGIS, audits, and program reviews.

The second major type of problem encountered in the data collection was where the definitions of the data elements in the alternative sources do not match those used on the FISAP. The effect of this is to dilute the power of the comparison of the two values, and hence the effectiveness of the analysis. Some of the lack of directly comparable definitions were anticipated. The discussion in Section 3 identifes the data elements, such as tuition and fees, that have comparable, but not exactly the same, definition. There were three unanticipated definitional problems.

One unanticipated problem with definitions occurred with the enrollment figures for nontraditional schools. The proposed alternative data source finally agreed upon was the Postsecondary Schools with Occupational Programs survey. Other data sources were investigated, in particular the Vocational Education Data System (VEDS). However, it was the opinion of the NCES staff responsible for the data that VEDS does not collect the kind of data that is contained in the FISAP. Moreover, it was felt that a number of schools in our sample would not be found in VEDS.

The Postsecondary Schools with Occupational Programs survey is a biennial survey of schools that asks for basic identifying information about the schools, much in the same way as the Institutional Characteristics file of HEGIS. The survey collects data on enrollment, cost of attendance, and other types of similar information. The problems with using this survey for the purpose of the 1982 FISAP analysis are first, that the relevant survey is for 1980-81, and second, the headcount enrollment figure from the survey cannot be directly compared because the FISAP asks for a breakdown of the number of new starts and continuing students by month. Initial attempts to compare the headcount enrollment from the survey and a constructed "average" enrollment from the FISAP proved to be unreliable. It is unknown whether the failure to adequately compare the values was due to the incomparability of the data definitions, the one-year difference between the data sources, or the inaccuracy of either the FISAP or the survey data.

The second unanticipated problem was the cross-year comparison of institutional expenditures. The plan was to compare level of institutional expenditures reported between the October 1981 and October 1982 FISAP. It was thought that the level of expenditures for the 1977-78 year would be reported on the October 1981 FISAP as it is on the October 1982 FISAP. This, in effect, would be a direct comparison since the amount for the 1977-78 year should not change between the two years. However, the October 1981 did not require the amount of institutional expenditures for the 1977-78 year, but rather the amount for the 1980-81 academic year.

The last unanticipated problem concerns the cross-year comparison of the maintenance of effort. The October 1981 FISAP asks for the level of institutional expenditures for 1981-82 for institutions that participated in CW-S or SEOG. The October 1981 FISAP required all institutions to submit three years of institutional expenditures. This means that the only valid comparison is between the 1981-82 expenditure figure and the last year of the three-year average, the 1980-81 expenditures.

The total effect of the problems encountered in the initial stage of data collection is to restrict the scope of the analysis. By far, the major problem was the missing document and/or missing data. For example, for 25 out of the 275 schools participating in the Campus-Based programs, no hard-copy FISAPs could be found. There could be many reasons for this including ED staff review of the FISAP without placing an "OUT" card in the folder, misfiled FISAPs, and so forth. Extraordinary efforts to track down these 25 were not employed in order to minimize staff disruption and because of the shortage of time to complete the data collection.

4.4 DATA COLLECTION AT THE INSTITUTION

Once the pre-visit data collection and analysis was completed, a comprehensive profile of the application data in the FISAP was constructed. This provides the starting point for the data collection and analysis using data gathered at the institution.

For each occurrence of an apparent error in the FISAP, a form was prepared for the institutional data collection. Figure 1 shows the form that was used to record

SUPPLEMENTAL FISAP QUESTIONS FOR

_____

As part of the Campus-Based Quality Control Study, Advanced Technology was asked by the Department of Education to compare the values on the application portion of the FISAP to alternative sources of information. We found what appear to be discrepancies in the following data items. Could you explain the reason for the apparent discrepancy and provide us with any supporting documentation.

Data Item	FISAP	Other Source
Tuition and fees		
Pell Expenditures		
Income grid		
- UG Dependent		
- Grad. Dependent		
- UG Independent		
- Grad. Independent		
UG Enrollment		
Grad. Enrollment		

Explanation #1 (_____) - _____

_____

_____

_____

Explanation #2 (_____) - _____

_____

_____

_____

Explanation #3 (_____) - _____

_____

_____

_____

FIGURE 3

INSTITUTIONAL DATA COLLECTION SHEET

information at the institution. Whenever there was an exception identified in the analysis, the FISAP value and the value from the alternative source were entered on the sheet along with the name of the alternative source. The data collectors were instructed during their training to ask the institution's financial aid administrator (FAA) about the apparent discrepancy, to explain how the discrepancy could have occurred, and what procedures the institution uses to generate the values for the FISAP.

The gathering of information about the discrepancies in the FISAP is an integral part of the overall data collection effort at the institution. The information can be obtained using a combination of two activities the data collectors will perform:

- Through the questions about the FISAP in the Institution Questionnaire
- Through requests for documentation for those data elements in the application.

Both of these collection procedures will yield information about the values of the items the institutions use in the application as well as the procedures and calculations used to produce the information that goes into the application. Examples of some returned institutional questionnaires appear in Attachment A.

The results of the institutional data collection will then be tied back to the pre-visit collection and analysis to give an in-depth look at the causes of errors in the application and what the magnitude of the impact is of those errors.

5.0 RESULTS OF ANALYSIS

This section reports the results of examining the data collected at ED. Since the analyses are comprised entirely of comparisons, there is no specific plan of analysis per se. Instead, in the first part of the section there will be a discussion of the general guidelines that were followed while performing the comparisons. The second part of this section will give an overview of the results of the analysis.

5.1 METHOD OF ANALYSIS

Once the data had been collected and entered onto the summary data sheet, the comparisons of values began. Section 3 identified the exact comparisons of data elements that can be made, and which comparisons are only approximate. The exact comparisons are self-explanatory from an analysis point of view. In the approximate comparisons, the method used becomes important.

A number of rules of thumb were used in the comparisons:

- For each data element for which there is an approximate comparison across two data sources for the same year, no error was reported if the values were within 10 percent of each other
- For each data element for which there is a cross-year comparison, no error was reported if the values were within 10 percent of each other
- An error was reported if there was a difference between the value on the FISAP data file and the hard copy FISAP
- If either the enrollment or tuition and fee revenues increased and the other decreased by a signficant amount, and error was reported
- If the enrollment increased and the total number of eligible aid applicants decreased significantly, an error was reported.

These rules were used as guidelines in doing the comparisons. In some cases, whether or not to report an error was purely a judgment call. The cross-year comparisons are particularly difficult because they are one-time snapshots of a difference. No trend analysis was developed that would make the evaluation of a change in value easier.

5.2 RESULTS

The in-house data collection uncovered numerous error on the application portion of the FISAP for a significant number of institutions in the sample. In total, 88 errors at 83 institutions were found. Figure 4 shows the frequency of the errors by type of error. Enrollment errors dominate the kind of errors made. Over two-thirds of all errors are enrollment errors. The next highest is tuition and fees error with twelve occurrences and then eligible applicant incomes with eight. The number of errors exceeds the number of schools because about 7 percent of the institutions made more than one error. The types of errors are self-explanatory except for the tuition and

DATA ELEMENT	FREQUENCY
Enrollment	60
Tuition and fees	6
Pell expenditures	7
Eligible Applicant Income	8
Institutional expenditures	1
Tuition and fees/enrollment	6

FIGURE 4

OCCURENCES OF ERROR ON THE APPLICATION PORTION OF THE FISAP

fees/enrollment error. This error occurred six times. This error is a combination of the change in tution and fees and enrollment. If, for instance, tuition and fee revenue increases substantially while enrollment drops, the school was flagged as being in error. It would be impossible to determine which one is in error so the combination was defined as an error.

The schools that made errors on the application come from all sectors of postsecondary education. Figure 5 shows the distribution of the 83 schools by type and control. Also presentd in Figure 5 is the distribution of schools who participated in the Campus-Based programs in 1983-84. As can be seen, it is not the proprietary schools who disproportionately make errors. Instead, the public institutions tend to make more than their share of errors. This is potentially disturbing given the fact that public schools receive such a large proportion of the funds. Attachment B shows the list of institutions that committed errors.

One interesting item of note from the frequency of errors is the relatively few number of institutions that committed more than one error. As mentioned above, only about 7 percent of the institutions had two or more errors. This tends to run counter to our hypothesis that there are some schools that make errors all the way through the application. Instead, most schools made one error. Figure 6 shows the six schools that made more than one error. This list excludes those institutions that committed the tuition and fees/enrollment error and only includes those with unrelated errors. Figure 6 shows that, like all the institutins making errors, those with multiple errors come from different sectors of postsecondary education.

The high frequency of enrollment errors could be due to one or both of the following reasons:

- The comparison was exact and there was no room for discretion over small differences
- The enrollment figures typically do not come from the financial aid office, but from the registrar or other administrative office, thus increasing the chance of error.

The Pell expenditures values are also virtually exact comparisons, but usually the amount of Pell expenditures is maintained by the financial aid office [redacted]ere is less chance for error. Moreover, the Pell amount can be taken directly from the Progress

TYPE AND CONTROL OF INSTITUTIONS	INSTITUTIONS WITH ERRORS	INSTITUTIONS IN CAMPUS-BASED PROGRAMS
Public University	26 (31%)	414 (10%)
Public 4-Year	9 (11)	93 (8)
Public 2-Year	19 (23)	800 (19)
Private University	10 (12)	587 (14)
Private 4-Year	14 (17)	699 (16)
Private 2-Year	2 (2)	198 (5)
Proprietary	3 (4)	1,473 (34)
TOTAL	83 (100%)	4,264 (100%)

FIGURE 5

INSTITUTIONS WITH ERRORS BY TYPE AND CONTROL

SCHOOL	ERRORS
St. Joseph's College (Private 4-Year)	Tuition and fees, enrollment
Rutgers (Public University)	Pell, enrollment
Soffolk Community College (Public 2-Year)	Pell, enrollment
Heidelberg College (Rublic 4-year)	Tuition and fees, enrollment
SUNY-New Paltz (Public University)	Pell, enrollment

FIGURE 6

SCHOOLS WITH MORE THAN ONE ERROR

Report while a number of the enrollment figures must be summed off of the HEGIS form.

Two points should be made concerning the Pell and enrollment comparisons. First, the Pell comparisons allowed a 1 percent tolerance before being declared in error. This allowance was made because many schools' expenditures do not match what the Pell Universe contains for either current authorization or net expenditures. This is due to the ongoing reconciliation process that continues for months after the award year. These adjustments are usually very minor and would account for the small discrepancies.

Second, many schools seem to have filled in their total enrollment figures for graduates and undergraduates directly off the HEGIS form. The total enrollment figures from the HEGIS include unclassified students. As a result, there really is no discrepancy that could be rectified by the institutional data collection. This problem is still an error. This error occurred in 46 schools. In addition, there were a significant number of schools for whom the original HEGIS enrollment matches the amount on the FISAP, but for some reason, the HEGIS enrollment figure was changed. The reasons for changes in HEGIS have to do with editing processes that also take place over several months. The impact of both of these discrepancies, unclassified students and changes on HEGIS, is minor as a percentage of total enrollment.

5.3 ANALYSIS PROBLEMS

The results section indicates a number of problems that were encountered in the analysis. The major ones are:

- Missing documents and data
- Lack of specificity on audits and program reviews
- Unexplained changes in HEGIS figures
- Documentation of changes between the hard copy FISAP and the FISAP data file.

The missing documents and data problems were discussed in the previous section. This problem is by far the most significant one restricting the analysis. Besides the

missing FISAPs, there were 27 instances where there was no HEGIS and 9 cases where an amount for Pell could not be found. As mentioned earlier, the definition of a reporting unit is different across types of reporting. Therefore, even though a particular school filed a FISAP, there is no certainty that, using that particular school's identification number, it will show up on either the Pell Universe or the HEGIS. In addition, many proprietary schools do not file a HEGIS. The frequency of missing HEGIS forms include only those schools that provided undergraduate and graduate enrollment on the FISAP, however, so there is no double-counting of missing forms.

The audits and any program reviews pertaining to the 1981-82 year revealed very few schools with problems on their FISAPs. Only five of the 275 schools had program reviews relating to activities in that year. Of those, there was little or no specific mention made of the application portion of the FISAP. Any problems that did occur had to do with accounting for expenditures of Campus-Based funds. In the case of audits, 15 out of the 275 were not found. In part, this was due to the microfiching process under way in the Audit Branch. Of the 260 audits reviewed, 10 audits were found to have FISAP exceptions. The majority of the audit exceptions had to do with the maintenance of effort figures. In one school, the enrollment figures (531 rather than 538 undergraduates), as well as the income grid of eligible aid applicants were found to be inaccurate. In two cases, the schools were under a general investigation for their administration of all student aid funds. In all cases, there was little or no specificity of what exactly was the problem.

In several cases there was a problem with changing values on the HEGIS form. When inspecting the HEGIS file, there were occurrences of HEGIS figures being manually changed from what was originally entered on the form by the institution. In most cases, the value on the FISAP matches the original value on the HEGIS. If there were changes made on the HEGIS, they were not reflected on the FISAP. This situation brings to light an analytical issue: Does an error occur if the FISAP value does not match the corrected HEGIS value? Since the editing of the HEGIS may be subsequent to the allocation of funds, is the institution in error? The reasons for changing HEGIS values are numerous but focus around cross-year edits, arithmetic errors, and so forth. The data collection for this analysis took into account arithmetic errors and entered only the correct value. However, there remained a number of unexplained changes of the HEGIS values.

The error reports that are returned to the institution for their review and correction do not always reflect all of the changes made to the FISAP file. There were seven occurrences of discrepancies between the FISAP data file value and the value on the FISAP form. It is unclear, at this point, however, whether the discrepancy arose out of the key entry error or from a submitted change from the institution. In all cases, there were no error reports containing the changed values.

6.0 CONCLUSIONS

The results of the in-house data collection showed that about 30 percent of the school in our sample committed one or more errors on the application portion of the FISAP. The majority of the errors are the enrollment figures that are supposed to come from the HEGIS Fall Enrollment survey. Contrary to a priori assumptions, the schools who committed errors are not concentrated in the proprietary sector, and schools do not tend to make many mistakes if they make one on the application. Instead, a disproportionate number of public schools make errors although all types of schools made errors. Also, very few schools made more than one error on the application, and of those who do, they make no more than two errors.

Perhaps more important than the numeric results are the discoveries made about the data sources used. Sections 4 and 5 have documented the problems encountered when doing the data collection and analysis. In particular, the following problems were found:

- Missing documents — FISAPs, audits, and program reviews
- Difficulty of identifying institutions — no common identifier other than name, differences in reporting units
- Accuracy of the data — all data seem to be continually updated resulting in possible differences in different versions of the same data file.

All three of these problems limit the accuracy and usefulness of this analysis because there may not exist a "best value" for a particular data element on the application. In a larger sense, the existence of these problems may be an indication that any attempt to validate application data with other data must start one step back. Instead of attacking the problem of how to validate the application, the other data must first be

cleaned up and then linked in some way to the Campus-Based System with a common identifier. Only then can one begin to validate data.

In terms of validating the data elements on the application, the following steps should be taken:

- **Restrict the current edits in the system to internal consistency and cross-year FISAP data.** The current edits are good, but they do not go far enough. Since the last year's FISAP is "clean" it should be used to evaluate the values on the current year's FISAP. This should be the first step of using other data files in the editing and updating process.

- **Begin to expand the use of other data files in cross-checking when they are determined to be "clean."** The Pell Universe file is a current file that is very accurate. It could be very useful for cross-checking. In the absence of a very "clean" file, tolerances could be established within the system to allow minor discrepancies to pass. In this analysis, a one percent tolerance was allowed for Pell. About 3 percent of the institutions still had errors that exceeded this tolerance. If an exact match had been used, about 10 percent of the institutions would have been flagged.

- **Work with the developers of other data files to determine the validity of their data for purposes of validation.** The editing cycles of a data file such as HEGIS may coincide with that of FISAP. If so, then an alternative needs to be found. It could be a cross-year HEGIS or something else.

- **Be creative with the kind of edits that are used.** In this analysis, the combination of tuition and fees and enrollment was used as an error measure. Others could be developed to check and cross-check the data in many different ways.

The number of errors in the application data requires that some steps for validating the data be planned. This analysis is a first step in that planning process because it has identified the kinds of problems one would meet in trying to use data for validation purposes.

ATTACHMENT A

SAMPLE FISAP DATA COLLECTION FORMS

INSTIT #128

SUPPLEMENTAL FISAP QUESTIONS FOR

Nicholls State Univ - LA

As part of the Campus-Based Quality Control Study, Advanced Technology was asked by the Department of Education to compare the values on the application portion of the FISAP to alternative sources of information. We found what appear to be discrepancies in the following data items. Could you explain the reason for the apparent discrepancy and provide us with any supporting documentation.

Data Item	FISAP	Other Source
Tuition and fees		
Pell Expenditures		
Income grid		
- UG Dependent		
- Grad. Dependent		
- UG Independent		
- Grad. Independent		
UG Enrollment	6097	5824 HEGIS
Grad. Enrollment	1129	1424

Explanation #1 (_____) - FISAP data was taken incorrectly from Fall 1982 Index Enrollment Statistics, Sept 16, 1982 instead of Fall 81 - " " " Sept 17, 1981. Report received per semester from Admissions

Explanation #2 (_____) - Fall 81 = 5824 / 1424 Fall 82 = 6097 / 1129 Hours have been ↑

Explanation #3 (_____) -

SUPPLEMENTAL FISAP QUESTIONS FOR

Skyline College - CA

As part of the Campus-Based Quality Control Study, Advanced Technology was asked by the Department of Education to compare the values on the application portion of the FISAP to alternative sources of information. We found what appear to be discrepancies in the following data items. Could you explain the reason for the apparent discrepancy and provide us with any supporting documentation.

Data Item	FISAP	Other Source
Tuition and fees		
Pell Expenditures		
Income grid		
- UG Dependent		
- Grad. Dependent		
- UG Independent		
- Grad. Independent		
UG Enrollment	8415	5775 (HEGIS)
Grad. Enrollment		

Explanation #1 (_____) - Information submitted to HEGIS indicated an evening enrollment and day enrollment. Only evening enrollment was picked up. Actual breakdown = Evening #5126 ; Day #2173 . Total = 7839 (computer printout attached that was sent to HEGIS.

Explanation #2 (_____) -

Explanation #3 (_____) :

SUPPLEMENTAL FISAP QUESTIONS FOR

Albany Jr. College - GA

As part of the Campus-Based Quality Control Study, Advanced Technology was asked by the Department of Education to compare the values on the application portion of the FISAP to alternative sources of information. We found what appear to be discrepancies in the following data items. Could you explain the reason for the apparent discrepancy and provide us with any supporting documentation.

Data Item	FISAP	Other Source
Tuition and fees		
Pell Expenditures		
Income grid		
- UG Dependent		
- Grad. Dependent		
- UG Independent		
- Grad. Independent		
UG Enrollment	1909 ✓	1827 (HEGIS)
Grad. Enrollment		

Explanation #1 (_____) - The FISAP reported figure is correct. The HEGIS reported figure is not at all familiar to the definition. We denied any knowledge or responsibility to it.

Explanation #2 (_____) -

Explanation #3 (_____) -

SUPPLEMENTAL FISAP QUESTIONS FOR

Indian River Comm College - FL

As part of the Campus-Based Quality Control Study, Advanced Technology was asked by the Department of Education to compare the values on the application portion of the FISAP to alternative sources of information. We found what appear to be discrepancies in the following data items. Could you explain the reason for the apparent discrepancy and provide us with any supporting documentation.

Data Item	FISAP	Other Source
Tuition and fees		
Pell Expenditures		
Income grid		
- UG Dependent		
- Grad. Dependent		
- UG Independent		
- Grad. Independent		
UG Enrollment	6047	3173 (HEGIS)
Grad. Enrollment		

Explanation #1 (_____) - FAA received data from the Asst Dean of Financial Affairs (Tommy Stein) & It is documented.

The Financial Affairs Office reports that

Explanation #2 (_____) - they sent an "FE-1" report to the Florida office of Education containing the same figure that the FAA used on "FISAPS" (6047). They do not know the origin or how the 3173 HEGIS was made.

Explanation #3 (_____) - They both feel the 6047 figure is correct.

SUPPLEMENTAL FISAP QUESTIONS FOR

Bucks. County Comm. College - PA

As part of the Campus-Based Quality Control Study, Advanced Technology was asked by the Department of Education to compare the values on the application portion of the FISAP to alternative sources of information. We found what appear to be discrepancies in the following data items. Could you explain the reason for the apparent discrepancy and provide us with any supporting documentation.

Data Item	FISAP	Other Source
Tuition and fees		
Pell Expenditures		
Income grid		
- UG Dependent		
- Grad. Dependent		
- UG Independent		
- Grad. Independent		
UG Enrollment	9708	9981 (HEGIS)
Grad. Enrollment		

Explanation #1 (_____) - Reports at School are matched - CAN'T EXPLAIN DISCREPENCIES, DOCUMENTATION SHOWS DIFFERENT NUMBERS AND SCHOOL COPIES OF HEGIS AND FISAP MATCH

Explanation #2 (_____) -

Explanation #3 (_____) -

ATTACHMENT B

SCHOOL	
George Wallace Community College	New England Conservatory of Music
Canada College	Newberry Junior College
Chaffey College	Wheelock College
Merritt College	Coppin State College
Skyline College	St. Joseph's College
University of San Diego	Michigan State University
Colorado State University	Inver Hills Community College
Southern Connecticut State College	Moorhead State College
Georgetown University	Southwest State College
Florida State University	University of Minnesota - Duluth
Indian River Community College	University of Minnesota - Morris
Miami-Dade Community College	Missouri Western State College
Palm Beach Community College	Delta State College
Albany Junior College	Mississippi College
Albany State College	Campbell University
Mercer University	Cleveland Technical Institute
Ball State University	North Carolina Central University
Indiana State University	Sacred Heart College
Valparaiso University	Franklin Pierce College
St. Mary's College	Atlantic Community College
Delta School of Business	Ramapo College of New Jersey
Louisiana College	Rutgers, State University of New Jersey
Louisiana State University A&M	Trenton State College
Nicholls State College	New Mexico State University
Bentley College	Bard College
Boston University	Boricua College
Emmanuel College	Daemen College
Fitchburg State College	Suffolk Community College
SUNY A&T	SUNY - New Paltz
SUNY - Stony Brook	Heidelberg College
Kent State University	Xavier University
Draughon's School of Business	Hillsdale Free Will College
Rogue Community College	Bucks Community College
Gannon University	Mercyhurst College
Pennsylvania State University	Shippensburg University

SCHOOL	
University of Pennsylvania	Anderson College
Furman University	Greenville Technical College
Rutledge College	University of South Carolina
Tennessee State University	Laredo Junior College
Texas State Technical College	Virginia Commonwealth University
Bennington College	University of Washington
University of Wisconsin - Milwaukee	