

## University Libraries as Economic Systems

*The possibility of developing an economic theory of libraries is explored. The concepts of economic theory, economic system, and economic model are discussed as potentially applicable to libraries. Two types of such models are developed from data drawn from university libraries. One predicts professional staff size from two variables: collection size and collection decentralization. The other identifies a set of library inputs composed of professional staff size, subprofessional staff size, and annual acquisitions rate as a consistently good predictor of library expenditures and a stable measure of library input.*

**T**HE FEASIBILITY OF AN ECONOMIC THEORY of the library has recently been suggested.<sup>1</sup> Were such a theory to be developed sufficiently to provide mathematical models, it is suggested that library planning and budgeting might be measurably improved.

Economic theory attempts to define and generalize the relationship existing among variables having to do with the production and distribution of wealth. Its method is largely deductive: on the basis of a set of known or assumed facts, a hypothesis is established and a model is set up. While the model may be simple or complex, it is a simplification of the real world to the extent that it does not include all the variables that could possibly be included. It is likely

to emphasize those that are important to the needs of the study and those that are readily measurable. The effect of those not included cumulates as error variance in the model. An example of an economic theory is that price tends to move to the level at which demand is equal to supply. Another is that a rise in price tends, sooner or later, to decrease demand and to increase supply.

Such theories require checking against data obtained from the real world. If empirical data support the relationship proposed by the theory, we may accept the theory as a useful assumption until additional evidence appears which conflicts with the theory, requiring its modification or abandonment. Because of the difficulties associated with environmental control in economic studies, economic theories indicate tendencies rather than universal laws.

Two different concepts are covered by the term "economic system." Both are concerned with the interrelationship of a set of economic variables. One is concerned with the ways in which a given society organizes its means of production and distribution of material wealth

---

*Maurice P. Marchant is acting director, School of Library and Information Sciences, Brigham Young University, Provo, Utah. This article was one of two winners of the 1975 Research Competition sponsored by the ALA Library Research Round Table and was presented by the author on June 30, 1975, at the Annual Conference of the association in San Francisco.*

and services and might be referred to as, for example, the American or British economic system. The other concept adheres to general systems theory and is concerned with identifying interrelationships between variables within a system: the processing of inputs into outputs, the effect of change in one variable on others, and so forth. This paper is concerned with the second concept, but is limited largely to the study of inputs.

The kinds of variables incorporated into economic models are those associated with production. They might be measures of wealth or productivity or those thought to affect or to be affected by them. Obviously, personnel, materiel, and money are important economic variables. For example, the model presented later in this paper in which physical decentralization serves as a predictor of staff size is an economic model because the emphasis is on personnel as a representation of funding. In another context, physical decentralization might be fitted into a political, rather than an economic, model in which the object is to study faculty influence.

In model building, theoretical assumptions of cause and effect relationships are made and tested. In doing so, within the concept of general systems theory, variables become categorized as input, intervening, and output variables. Speaking generally, they might be described as follows. Inputs are those things that enter the system from outside, such as books and personnel in a library. Outputs are products which are created in the system with the use of inputs and exported back into the environment, such as library service. Intervening variables are affected by inputs and, in turn, affect outputs. But the label given a variable under one set of circumstances may not apply in another. Thus an input variable in one system might be an intervening or output variable in another system.

Variables are also categorized as inde-

pendent and dependent. An independent variable has the capacity to predict the value of a dependent variable, and there is often a presumption of causal relations between them. These terms are more general than, but not unrelated to, the concepts of input, intervening, and output. Within a systems study, one would expect inputs to be independent variables, outputs to be dependent variables, and intervening variables to be both.

Two relevant research projects regarding libraries have emerged recently. One computed the annual growth rates of several variables in academic libraries and developed prediction equations for estimating various staff, collection, and cost values.<sup>2</sup> The other, noting that the distribution of many library statistics is skewed, approached the study of those statistics with improved success through their logarithmic values.<sup>3</sup>

These studies provide a sense of confidence to the assumption that generalized influences are at work which affect all or large groups of academic libraries. If so, their identification and measurement may help in the construction of an economic theory of the library which librarians might find useful in decision making.

The thrust of this paper is to provide recently identified evidence supportive of that position. However, these studies have been confined to academic libraries of limited size range, above 500,000 volumes. They were undertaken to explore a set of readily available data for evidence that might support the concept that libraries are economic systems and to identify points of departure for further research. Two groupings of predictive models will be discussed, one dealing with the size of the professional staff and the other with measures of input.

#### PROFESSIONAL STAFF SIZE

It was found in a study of twenty-two libraries of Association of Research

Libraries (ARL) member institutions that the ratio of professional staff members (mostly librarians) to 1,000 students was 4.4 (with a standard deviation of 3.5).<sup>4</sup> Comparisons with other variables in the study demonstrated that the ratio had a high relationship with several funding measurements. It took its place among them as a measure of library wealth, indicating that it reflects financial input in relation to the number of students who have a potential call on the library. This relationship was confirmed by factor analysis as well as by the magnitude of the Pearson product-moment correlation coefficients, which was as high as .99.<sup>5</sup> It was apparent that funding is a fairly good predictor of professional staff size. The high standard deviation relative to the mean of the librarian-to-student ratio results from both high variance in wealth among libraries and a skewing toward high wealth caused by a few especially wealthy libraries.

The number of professional staff members was also found to be highly related to several other organizational variables. Moreover, when their interrelationships were graphically depicted, number of librarians occupied a central position among them much like the axle of a wheel.<sup>6</sup> These variables, along with their correlations with professional staff size, were: (1) collection size, .72; (2) currently received serial titles, .54; (3) number of volumes acquired during the school year, .56; (4) physical decentralization of the collection, .71; and (5) number of doctoral degrees (excluding law, medical, and dental doctorates) granted that year by the university to which the library belongs, .62.

It is not difficult to imagine staff size being affected by these variables. The size and growth rate of the collection generate work to process and service it. A large graduate program, which is the primary justification of a large collection, would be expected to generate service demands. Overfragmentation of

the collection could cut down on the efficient use of personnel.

However, considering these variables' lack of independence from each other and the small number of libraries in the study, it would be surprising if they could all fit significantly into a formula predicting professional staff size. The easiest procedure for constructing an optimum predictive equation is through linear multiple regression analysis, in which combinations of independent variables are tested for significance and compared for predictive capacity. The best combination, as it turned out, included two variables: collection size and decentralization. Together, they explain almost 80 percent of the variance in professional staff size. If these relationships are causal and accurate, they offer help in predicting staff needs under changing conditions of collection size and decentralization. The predictive equation is

$$Y = 22.9 + 0.235X_1 + 67.8X_2$$

in which  $Y$  is number of professional staff members,  $X_1$  is collection size in 10,000s of volumes, and  $X_2$  is the decentralization index. The decentralization index is computed from the formula

$$D = B^2/C_t$$

in which  $D$  is the decentralization index and is equal in this case to  $X_2$ ,  $B$  is the number of branches, and  $C_t$  is the total number of volumes (in 1,000s) in the university library system and is ten times the value of  $X_1$  in this case. The derivation of the formula for decentralization is explained in the dissertation.<sup>7</sup> While the correlation between the independent variables was insignificant (.28), the involvement of collection size in both made the formula nonlinear.

The prediction equation has certain limitations. First, it was derived from libraries varying in size from 500,000 to 2,100,000 volumes and with a mean size of 1,160,000 volumes. The mean of the decentralization indexes was 0.084, and



the indexes ranged from 0.000 to 0.586. The equation functions best with libraries having values close to the mean, and the extent of error can be expected to increase as the values deviate from it.

There is also a certain potential for error in the prediction due to the 20 percent of the variance which was unaccounted for. The expected error is as much as 17 about once in twenty cases.

#### SETS OF LIBRARY INPUTS

The usefulness of professional staff size as a measure of input requires testing. But it is part of a more general question which asks what measurements constitute interrelated sets of inputs. It would be helpful if a set of mutually complementary inputs could be identified which are stable over time.

Inputs in this case consist of resources entering the library from the environment. Funding can be thought of as an input. So can the resources the budget provides, such as personnel and library collection components.

The dollar has been used as a measure of input by most libraries; but it has both advantages and disadvantages. As a means of exchange, it can stand for many different inputs, including personnel playing different roles and drawing differing salaries, books and serials, and various other materials and services libraries need. Consequently, the dollar can be used as a unit of input by which, in one sense, various inputs can be compared. For example, a librarian costing \$10,000 a year can be equated to 1,000 books averaging \$10 each.

But the potentially available freedom to choose what inputs to exchange the library's budget for may be delusionary. If a given set of goals is desired, its actualization may predetermine the optimum mix of inputs required. If so, and we understand the optimum system that will achieve it, the choice of inputs and their quantities have already been deter-

mined, and the inputs are not independent of each other.

Another weakness in using money as a measure of input is its instability during periods of inflation or recession. Consequently, it would be helpful if other, more stable, inputs could be identified.

#### RESEARCH METHODOLOGY

To study input stability, two sets of data were subjected to regression analysis. The first had been collected for the dissertation. It included three measurements of funding (total operating expenditures, staff expenditures, and library materials expenditures) and five measurements of the basic groups of resources funded by libraries (the total number of staff members, professional staff size, subprofessional staff size, the number of volumes acquired, and the number of current serials received). However, it was also limited to one year's data from only twenty-two libraries. The second set was the data based on punched cards regarding the libraries of fifty-eight ARL member institutions compiled by the Purdue University Library and Audio Visual Center.<sup>8</sup> While it lacked measurements for number of current serials received, the sample size was much larger and the data covered twenty-one years. Consequently, the smaller data base was used for a series of preliminary analyses to determine whether the lack of that one measurement might be a serious loss. The second set was then analyzed, first, to see to what extent it confirmed the preliminary analyses and, second, to determine the extent to which the relationships might have varied over time.

#### PRELIMINARY ANALYSIS SERIES

Total operating expenditures were best predicted, using the first set of data, by the size of the professional staff and number of acquisitions. Once they had entered the regression analysis,

none of the other measurements had a significant further predictive capability. The two, as a set, accounted for 83 percent of the variance in total operating expenditures.

Expenditures for library materials were best predicted by annual number of acquisitions and the professional staff size, which accounted for 76 percent of the variance in library materials expenditures.

Staff expenditures (including salaries and wages) were best predicted by the size of the professional staff alone, which accounted for 77 percent of the variance in these expenditures. None of the other potential dependent variables, including the number of subprofessional staff members, contributed significantly to the prediction.

The relationships between the three expenditures were also probed. Adding the staff and library materials expenditures together accounted for nearly 90 percent of the total operating expenses in the average of the libraries studied and would be expected, therefore, to provide an excellent prediction. Staff expenditures alone explained 91 percent of the variance in total expenditures, and library materials expenditures added 8 percent more to the explanation, for a total control of 99 percent.

Anticipating that the relationships identified in these preliminary analyses would be generally consistent with those in the larger study, several insights were possible.

First, the number of current serial titles offered little supplementary predictive potential to any of the expenditure variables. This is not to say that the cost of serials is unimportant. Rather, it suggests that (1) the number of serial titles is not a good indication of serial cost, and (2) the number of serial titles varies with such other variables as number of professional librarians to such an extent that its predictive potential, such as it is, is largely duplicative.

Second, the strongest predictors of total operating expenditures are professional staff size and number of volumes acquired. Of these, professional staff size appears to be the stronger. Number of volumes acquired affects expenditures largely through library materials expenditures, as would be expected; but professional staff size is an important predictor for both library materials and staff expenditures.

Third, with an increase in the number of libraries in the study, number of subprofessional staff members might emerge as a significant predictor of total operating and staff expenditures.

Fourth, total staff size did not provide a useful index for this study. The components of it, professional and subprofessional staff size, are more useful.

#### ANALYSIS OF SERIES OVER TIME

The preliminary analyses provided a sense of confidence that the lack of data regarding serials from the larger data base would not result in the lack of an important variable for the purpose of this study. As the regression analyses were completed, using the data for each year for making a set of regression analyses like those in the preliminary series, the patterns that had been expected emerged. Since there was some overlap between the libraries represented in the two data bases, this should not come as a surprise. But neither was it certain beforehand. Of the twenty-two libraries in the preliminary series, four were not among the fifty-eight libraries of the larger series. No library over 2,100,000 volumes in size in 1968 was among the twenty-two in the preliminary series, whereas twelve of the fifty-eight in the larger data base were above that size in 1968. That the pattern was similar between the two samples and from year to year in the second sample suggests that the pattern can be relied upon, within limits, from year to year and across a fairly wide

size range. Moreover, predictions of individual library expenditures and their confidence intervals were made possible.

It is that measure of consistency, supplemented by the Baumol and Marcus and Pratt observations, which provides encouragement to persist in the search for a theory of library economy.

One purpose of this study was to try to determine whether the staff size and acquisitions measurements might have greater stability over time than the expenditures measurements. One way to test stability is to see how constant the cumulative proportionate variance, as a measure of predictive capacity, remains from year to year as compared with the regression coefficients which are generated in sets of yearly regression analyses in which the staff size and acquisitions variables serve to predict the expenditures levels. Another evidence is the extent to which variance in the cumulative proportionate variance and regression coefficients is a function of time.

If the ability of a set of independent variables to predict the value of a dependent variable remains constant, the cumulative proportionate variance would not change. But the effect of inflation on the cost of the independent variables over the years would cause changes in the regression coefficients. The magnitude of variance from year to year is measured by the standard deviation. The extent to which the variance is constant over time would be expressed by a simple correlation between value of the variable and the year.

Summaries of the sets of analyses are given in Tables 1 through 4. They provide the data required to evaluate (1) the ability of specific sets of independent variables to predict library expenditures, (2) variance over time in that predictive capacity, (3) the mean value of regression coefficients associated with each independent variable in the set, (4) variance in the regression coefficients, and (5) the extent to which variance was a function of time. In addition, in order that a rough comparison can be made of the difference in variance between the regression coefficients and cumulative proportionate variance, the standard deviations were normalized by showing them as a ratio of their mean values. This normalized value is known as a coefficient of variation.

Table 1 summarizes data regarding the relationships of total operating expenditures with staff and library materials expenditures during the twenty-one-year period. These data are presented largely to provide a basis for comparison with the other three tables. This set of independent variables is shown to be an excellent predictor of total operating expenditures (predicting, in an average year, with 98.47 percent accuracy) with little deviation in predictive capacity from year to year (1.94 percent standard deviation). Staff expenditures is the more stable predictor. (Comparing the coefficients of variation determines their relative variability.) Neither the regression coefficients nor the cumulative proportionate variance changed in

TABLE 1

SUMMARY OF ANALYSES RELATING TOTAL OPERATING EXPENDITURES TO STAFF AND LIBRARY MATERIALS EXPENDITURES OVER TWENTY-ONE YEARS

	Mean	Standard Deviation	S.D./Mean	Correlation with Year
Regression coefficients				
Staff expenditures	1.1553	0.1281	0.1114	-0.1390
Library materials expenditures	1.1085	0.2383	0.2150	-0.1951
Cumulative proportionate variance	0.9847	0.0194	0.0197	-0.1089

TABLE 2

SUMMARY OF ANALYSES RELATING TOTAL OPERATING EXPENDITURES TO PROFESSIONAL AND SUB-PROFESSIONAL STAFF SIZE AND NUMBER OF VOLUMES ACQUIRED OVER TWENTY-ONE YEARS

	Mean	Standard Deviation	S.D./Mean	Correlation with Year
Regression coefficients				
Professional staff size	12418.	6028.	0.4854	0.9179
Subprofessional staff size	3915.	1601.	0.4088	0.4658
Acquisitions	4.0917	2.1690	0.5301	0.2736
Cumulative proportionate variance	0.9309	0.0309	0.0385	-0.2194

a pattern associated with time. (The correlations with year were low and not significant at the .05 level, which was the lowest significance level tested throughout the study.) Since all three variables were affected by inflation, these nonsignificant correlations were expected.

Regression coefficients associated with more constant inputs as predictors of expenditures, being immune to inflation, should be more highly correlated with time because the resources acquired will increase at a slower rate than the funding expended to acquire them. Tables 2, 3, and 4 provide the data generated to test the constancy of three such input variables and their predictive capacity. In each case, the input variables listed are those which contributed significantly, at the .05 level or better, to the prediction and which, as a set, provided the best prediction available.

All three input variables were able to enter the analyses, in Table 2, predicting total operating expenditures. Note that the larger sample size and breadth of years covered allowed for the inclu-

sion here of subprofessional staff size, which did not emerge in the preliminary study. In seventeen of the twenty-one yearly analyses, it emerged as a better primary supplement to professional staff size than did number of volumes acquired.

Together they explained an average of 93.09 percent of the variance with a standard deviation of 3.09 percent, a substantial improvement over the preliminary analysis results of 83 percent. The inclusion of the larger number of libraries appears to have improved the statistical measurements of the relationships.

While the prediction is lower than was attributed to the two expenditures variables in Table 1, it is remarkably high and stable. In only one year did it drop below 90 percent.

The coefficients of variation (S.D./Mean) of the regression coefficients were much higher than those in Table 1 and indicate a fairly high variation in regression coefficients over the years. Much of the variation was not random but, rather, was the result of increases

TABLE 3

SUMMARY OF ANALYSES RELATING LIBRARY MATERIALS EXPENDITURES TO PROFESSIONAL STAFF SIZE AND NUMBER OF VOLUMES ACQUIRED OVER TWENTY-ONE YEARS

	Mean	Standard Deviation	S.D./Mean	Correlation with Year
Regression coefficients				
Professional staff size	3491.	1691.	0.4844	0.8186
Acquisitions	2.7042	1.1832	0.4375	0.5000
Cumulative proportionate variance	0.7508	0.1441	0.1919	-0.5109



TABLE 4

SUMMARY OF ANALYSES RELATING STAFF EXPENDITURES TO PROFESSIONAL AND SUBPROFESSIONAL STAFF SIZE OVER TWENTY-ONE YEARS

	Mean	Standard Deviation	S.D./Mean	Correlation with Year
Regression coefficients				
Professional staff size	8845.	4549.	0.5143	0.9198
Subprofessional staff size	3119.	955.	0.3062	0.6896
Cumulative proportionate variance	0.9371	0.0495	0.0528	-0.3576

over the years that are the result of inflation. That was particularly true of the regression coefficients associated with professional staff size and, to a more moderate extent, with subprofessional staff size.

Tables 3 and 4 summarize the data from the sets of analyses predicting library materials expenditures and staff expenditures, respectively. As expected from the preliminary analysis, professional staff size entered both sets. In addition, subprofessional staff size entered the staff expenditures set, and acquisitions entered the library materials set.

In each set, the correlations between year and regression coefficients for both input variables were positive and significant, indicating that the regression coefficients grow larger over the years. The staff size variables explained 93.71 percent of the variance in staff expenditures, which is comparable with the predictive power of the three input variables in Table 2. But control over variance in library materials expenditures, of 75.08 percent, was less strong and was associated with a standard deviation more than four times as large as was found in total operating expenditures.

The correlations of year with the cumulative proportionate variance associated with staff and library materials expenditures indicate a tendency for a decline in predictive power in recent years. The decline is more pronounced regarding library materials expenditures, with which the correlation is sig-

nificant at the .05 level, than regarding staff expenditures.

Overall, the data support the hypothesis that the staff size and acquisitions input variables have retained their powers to predict library expenditures. But their control has been better over total operating and staff expenditures than over library materials expenditures. Apparently, stability of prediction is greater for more general funding measures than for more specific ones.

#### CONCLUSIONS

A pattern of relationship between major personnel and materials inputs with library expenditures has persisted in American university libraries for two decades. The major inputs are professional and subprofessional staff size and number of volumes acquired. Together, they constitute a set of inputs that are stable over time and can be used in systems studies comparing conditions in one year with those in another in place of the expenditures variables which are not stable.

While a budget officer may not be willing to predict next year's funding needs in a specific library without greater specificity than these three gross measurements provide, that officer could probably estimate this year's expenditures within a margin of error as small as 6 percent with 95 percent confidence and as small as 7.5 percent with 99 percent confidence. With some experience, he or she could likely improve substantially.



Such accuracy is possible because many other expenditures are reflected in these three. Materials, such as paper, pencils, and card stock, are a function of the number of personnel and books acquired. While an administrator might control the costs involved in such items somewhat by bulk purchases and watching for bargains, control through these types of activities is only marginal. Moreover, attempts to economize unduly in these areas can readily result in operational bottlenecks that inhibit effective service. When a library acquires a new reference assistant, the resource commitment that will result, after the library has adjusted to the stress imposed by that increase, will be a great deal more than that staff member's salary and fringe benefits. It is more likely to be nearer \$20,000 than a \$10,000 annual salary.

The demonstration, in the first part of this paper, that professional staff

size is a function of decentralization and collection size is a minor contribution to a developing economic theory of libraries. It suggests that organizational patterns might be a fruitful area for future study into that development. Another area would be to identify the interrelationships between the three staff and acquisitions input variables and other resources, perhaps as part of the type of systems analysis studies and model building carried out with the use of techniques and concepts used in econometrics.

Radical changes in the pattern of library organizational behavior might affect the basic economic patterns identified in this study. But the pattern has remained fairly constant over a twenty-year span characterized by constant change. At least for the near future, those relationships are likely to persist and to provide a basis from which future costs can be predicted.

#### REFERENCES

1. Michael K. Buckland, "Toward an Economic Theory of the Library," in Robert S. Taylor, ed., *Economics of Information Dissemination: A Symposium* (Frontiers of Librarianship 16 [Syracuse, N.Y.: School of Library Science, Syracuse University, 1974]), p.68-80.
2. William J. Baumol and Matityahu Marcus, *Economics of Academic Libraries* (Washington: American Council on Education, 1973).
3. Allan D. Pratt, "A Theory of Lognormal Size Distribution of Academic Libraries in the United States" (Ph.D. dissertation, Univ. of Pittsburgh, 1974).
4. Maurice P. Marchant, "The Effects of the Decision Making Process and Related Organizational Factors on Alternative Measures of Performance in University Libraries" (Ph.D. dissertation, Univ. of Michigan, 1970), p.108-11.
5. *Ibid.*, p.120-21.
6. *Ibid.*, p.123.
7. *Ibid.*, p.96-97.
8. Oliver C. Dunn and others, *The Past and Likely Future of 58 Research Libraries, 1951-1980: A Statistical Study of Growth and Change* (8th issue; Lafayette, Ind.: University Libraries and Audio Visual Center, Purdue University, 1972), p.5.