R. MARVIN McINNIS

# The Formula Approach to Library Size: An Empirical Study of Its Efficacy in Evaluating Research Libraries

*Formula approaches to the determination of adequate library size, especially along lines developed by Clapp and Jordan have in recent years become a significant element of the librarians' arsenal. Nevertheless, as has all too often been pointed out by those responsible for budgets and funding, the empirical basis of the Clapp-Jordan formula is rather vague. In this paper the possibilities of using statistical regression analysis to provide such an empirical analysis are reviewed. The results indicate that it is indeed difficult to provide such an empirical foundation but it can be shown that there is not likely any upward bias to the Clapp-Jordan formula.*

THE UNIVERSITY LIBRARIAN is continually faced with a need to answer the question of whether his library collection is large enough to support the teaching and research activities of the university in anything like an adequate way. One device that he has at hand, and one that he has been turning to quite frequently in recent years, is a formula which is based on the careful judgment of experts in library evaluation, and which is intended to indicate a minimum scale of adequacy. Several such formulae are available although most of these stem from the one suggested by Clapp and Jordan.[1] Clapp and Jordan begin by listing the variables that should be relevant to the determination of the size of collections for academic libraries.

1. the size and characteristics of the student body
2. the size and research commitment of the faculty
3. curriculum—numbers of departments, courses, etc.
4. methods of instruction
5. the availability of study places on campus
6. proximity to other libraries
7. the intellectual climate of the university.

In their proposed formula, Clapp and Jordan took into account only the first three of the above variables. The remaining ones are admittedly difficult to conceive of in a quantitatively measurable fashion. In brief, the formula proposed by Clapp and Jordan states: Begin with a basic library with a number of volumes indicated by one or another of the well-known select undergraduate libraries. Then add 100 volumes per full-time faculty member, 12 volumes per enrolled student and 12 additional volumes per undergraduate honors student, 335 volumes per major undergraduate subject offered, 3,050 volumes per MA field offered, and 24,500 volumes for every field in which study for the PhD is undertaken at the institution. What one arrives at is an indication of the minimum number of volumes that the library should have if it is to

*Mr. McInnis is associate professor of economics, Queen's University, Kingston, Ontario, Canada.*

perform at all adequately as an academic library. Similar formulae are offered for current periodical titles and for government documents. The weights in the formulae are judgmental although based on a number of indicators of good library practice and the magnitude of bibliographic materials relevant to study and research at the university level.

What is rather surprising is that one can find so little in the way of reported attempts to determine empirically whether the Clapp-Jordan formula or any variant of it generally fits existing academic libraries. In particular, one would like to know whether the weights in the formula have a reasonably accurate empirical basis. It would seem that this problem would be suitable for solution by linear regression analysis.[2] The Clapp-Jordan formula can be written as a weighted sum of several variables, all of which are quantitatively measurable. Let us adopt the following symbols for the variables:

$a_0$ = a constant representing a minimum viable undergraduate library (Clapp-Jordan say 50,750 vols.)
F = the number of faculty
E = total number of students enrolled
H = number of undergraduate honors students
U = number of major undergraduate subjects
M = master's fields offered
D = doctoral fields offered

The Clapp-Jordan formula, for the number of volumes, V, can then be written:

$$V = 50{,}750 + 100F + 12E + 12H + 335U + 3050M + 24{,}500D \qquad [1]$$

The issue of concern here is the weights applied to the variables in [1]. Let us take a step back and treat them as unknowns. Then with statistics drawn from a sample of universities for each of the variables in the equation one could use regression analysis to estimate the values of the terms in equation [2]:

$$V = a_0 + a_1F + a_2E + a_3H + a_4U + a_5M + a_6D + e \qquad [2]$$

The variable e represents a random error. That is to say, equation [2] will not give a precise prediction of V but one that will be in error to some degree. The estimates are made on the assumption that these errors are randomly distributed.[3] A further, and very important assumption for what follows, is that each of the variables on the right-hand side of the equation exerts an independent influence on V.

An empirically estimated formula is such a natural extension of the Clapp-Jordan approach that it is a bit surprising that there is so little record of experimentation along these lines. I have found only one report of work of this type. Edwin W. Reichard and Thomas J. Orsagh used regression analysis to account for library expenditures and, what is more pertinent here, holdings of the libraries of colleges and universities.[4] They studied random samples of about three hundred institutions for the two years 1952 and 1962. However, they related numbers of volumes only to variables indicating the size of the institutions and made no attempt to take into account breadth and diversity of programs. In that sense they make only a partial study of the formula approach. Their size variables relate to faculty and students, as do those of the Clapp-Jordan formula, except that Reichard and Orsagh separate graduate from undergraduate students. Equation [3] restates the results they obtained by regressing the number of volumes on the number of undergraduate students ($E_u$), the number of graduate students ($E_g$), and the number of faculty (F) in 1962.[5]

$$V = 27{,}100 - 9.6E_u - 59E_g + 969F + e \qquad [3]$$
$$R^2 = .75$$

At first glance these results would appear to be a distressing commentary on the Clapp-Jordan formula. The coefficients bear no resemblance at all to those proposed by Clapp and Jordan. In equation [3] the preponderant influence upon library size is the number of faculty, and Reichard and Orsagh emphasize this as a particularly striking attribute of their findings. Numbers of students, both undergraduate and graduate, appear to be negatively related to library size. Moreover, the constant term is much smaller than the basic academic library with which Clapp and Jordan claim one should start.

The last of these features of their results may just reflect the fact that the data used by Reichard and Orsagh were drawn from a list of academic institutions that must exhibit varying degrees of adequacy in their library facilities. If the sample is truly random it will contain inadequate as well as adequate libraries. More seriously for our present purposes, however, is a flaw in their analysis which effectively invalidates the conclusions they reach. The size variables $E_u$, $E_g$, and F are patently not independent of each other. Indeed, one would expect them to be rather highly correlated. One would probably get quite similar results using either faculty or student enrollment separately but the statistical analysis has no way of isolating their true separate influences. The fact that in this particular application F came out with a positive sign and a large coefficient and the E variables did not warrants no conclusion at all. The technical term given by statisticians to this problem is *multicollinearity*.

I want now to report on an effort at a more valid approximation to the Clapp-Jordan formula by purely empirical means. It differs from the work described above in three respects. First, I attempt to rid the analysis of multicollinearity, although without complete success. Second, I incorporate a measure of the diversity of academic programs at the universities included in the analysis. Third, I focus only on established graduate schools. This last marks a sharp deviation from either Reichard and Orsagh or Clapp and Jordan. The rationale is primarily that the present study is just a part of a larger one that has as its aim the development of an alternative approach to evaluating the adequacy of library resources for graduate training and research. What I have done here could be repeated with reference to a broader set of institutions with only a modest effort. However, the present results may have some general interest beyond the particular question that concerns me—the adequacy of libraries for graduate study and research. From that point of view I have selected as a frame of reference the successful graduate schools of the United States. Operationally, I define these to be the thirty-six leading universities in the United States in terms of PhD's granted during the period 1959–62.[6] There are other graduate schools, of course, but those incorporated in the analysis are the ones that appear to have been clearly successful. There is some presumption that their library resources meet a standard of adequacy (or at least if they do not it has not impaired their viability as graduate schools to any noticeable degree).

What I am dealing with is not really a sample but the whole population of most successful graduate schools (at least on the pragmatic definition of success that I have given). Statistically, however, this might be thought of as a sample of the population of conceptually successful graduate schools. Looking at the evidence in this way it is difficult for us to judge how random the sample may be. Since a good number of the institutions far exceed the minimum levels of adequacy postulated by Clapp and Jordan, we might fairly expect the predictions of library size given by the formula estimated here to overshoot the results obtained with the Clapp-Jordan formula.

Numbers of volumes and periodical titles were obtained from statistics published by the American Library Association.[7] Enrollment, both graduate and undergraduate, and numbers of faculty were from publications of the U.S. Department of Health, Education and Welfare.[8] These bodies of data are reasonably well known and whatever weaknesses they may have are not peculiar to the present study. The number of fields in which institutions offer study for the PhD degree, the measure of diversity of programs utilized here, was developed from listings in *American Colleges and Universities*.[9]

To get around the problem raised by the correlation between numbers of students of both sorts and the number of faculty members, one of these variables had to be chosen as the primary indicator of size of the institution. Without intending to enter into the ideological debate on the appropriate locus of power on campuses, I adopted the number of faculty as the primary indicator of size. I am dealing with major research institutions where it is likely that this indicator of the size of the institution would have the most bearing on the determination of library size. The number of students is then introduced relative to the number of faculty and the number of graduate students relative to the total number of students. The important point here is the introduction of the various measures of size as a group with some effort to capture the independent influences of each variable. It would have been just as suitable to use enrollment as the primary measure of size and to add variables for the number of faculty members per hundred students.

Numbers of fields is suggested by Clapp and Jordan as an indicator of the breadth of program offered by an institution. I make no attempt to handle MA fields separately. Mostly they will be fields in which the PhD is offered and, certainly, for the selection of universities being studied, the addition of fields where the MA but not the PhD is offered would not likely add much to the analysis. The definition and identification of fields of doctoral study is both difficult and ambiguous. The lists of fields offered by institutions usually has an administrative basis and often reflects peculiarities of the historical development of the institution. Moreover, fields appear to be more narrowly defined in the natural sciences than in the social sciences. Language fields offer an acute example of variations in the designation of offerings. Some universities note very specifically what languages they offer (Spanish, Italian, French, Portuguese) whereas others organize them into broad groupings (Romance Languages). There is no real alternative to grouping specific offerings into those broad categories that appear to be fairly commonly used, so as to assure a tolerable level of uniformity among universities.[10] Area studies proved to be much more difficult. Since the universities generally regard the comprehensive study of a particular area as a distinct field, I accept this assumption. It is really doubtful if the numbers of fields can be given any very precise interpretation, given the possible variations in the treatment of field designations. At any rate, an attempt has been made here to specify a standard list of fields and to note which of them are offered at each university. However, the ambiguities in the basic evidence are such that it should not be thought that the data used here conform precisely to such a neat tabulation. Little more can be said than to caution readers about the weakness of this variable.

Science and nonscience fields have been tallied separately on the grounds that their bibliographic needs may differ, especially when summarized in such a gross way as numbers of volumes or periodical titles. The variables used in this analysis are listed below.

$V_1$ = library holdings in volumes (in '000's)
$V_2$ = current periodical titles (in '000's)
F = number of faculty[11]
S = student enrollment per 100 faculty members
G = graduate students per 1,000 students enrolled
D = total number of doctoral fields
$D_s$ = number of natural science doctoral fields
$D_n$ = number of nonscience doctoral fields.

Several regression equations were estimated, utilizing various combinations of the above variables. The specification of the relationship which, a priori, appeared to be most promising gave the results shown in equations [4] and [5]. The numbers in parentheses below the coefficients of the explanatory variables are values of the statistic t that is used in evaluating the statistical significance of the estimated coefficients.

$$[4] \quad \begin{aligned} V_1 = -875.30 &+ \underset{(.67)}{.089F} + \underset{(.23)}{.007S} \\ &+ \underset{(.01)}{2.504G} + \underset{(.77)}{23.336D_s} + \\ &\underset{(.50)\ (2.26)}{97.980D_n} + e \end{aligned}$$

$\bar{R}^2 = .29 \qquad F = 3.924$

$$[5] \quad \begin{aligned} V_2 = -17.450 &+ \underset{(1.92)}{.003F} + \underset{(1.18)}{.006S} \\ &+ \underset{(1.34)}{.037G} + \underset{(1.64)}{.101D_s} + \\ &\underset{(.31)\ (2.94)}{.891D_n} + e \end{aligned}$$

$\bar{R}^2 = .54 \qquad F = 9.216$

The result for numbers of periodical titles, equation [5], is much stronger than for numbers of volumes of books. The main conclusion that can be reached from both equations, however, is that the explanatory variables that are tested do not perform especially well. The random element is large, particularly in the case of $V_1$. The proportion of the variance of $V_1$ that is accounted for by the indicators of size and diversity is only .29.[12] For $V_2$ the regression equation does a little better and $\bar{R}^2$ is .54. That is to say that the equation is able to account for 54 percent of the variation in numbers of periodical titles. In both cases, though, the random element is substantial. Cross-section regressions that have genuine explanatory content not infrequently have a low $\bar{R}^2$. However, we cannot be entirely pleased with results that, at best, account for only half of the variations in library size among institutions. Even more serious is the failure of most of the explanatory variables to show up with a statistically significant influence on library size. If we accept an approximate test that the value of the t statistic should exceed two before we conclude, with a probability of .95, that any coefficient probably exceeds zero, we find that only one of the postulated variables has a significant influence upon library size. That is $D_n$, the number of nonscience fields offered. That one variable accounts for almost all of the explanatory power of the regression equation. For numbers of volumes, the coefficients of all other variables are small and have such large variances that one could not conclude that they are really different from zero. It is doubtful that it is even worth pointing out that the coefficients imply very different weights from those of the Clapp-Jordan formula. Even in the case of the one statistically significant variable, the regression results imply that a library adds about 98,000 volumes per nonscience field in which it offers the PhD. That is several times the figure used by Clapp and Jordan. The result would also imply that none of the other variables —faculty or students or science doctoral programs—likely has any significant independent influence on library size. Such a result, although apparently damaging to the Clapp-Jordan formula, lacks credibility.

A major difficulty lies in the nasty matter of multicollinearity which, in spite of the care which I tried to exercise, has not been excluded. It turns out that the number of doctoral programs is rather highly correlated with the number of members of faculty. The simple correlation coefficient with total doctoral programs is .75. With the number of nonscience doctoral programs it is .66. In either case the correlation is far too high to provide for separate estimation of the influence of numbers of faculty and numbers of programs.

The implication of the statistical analysis should be clearly put—the Clapp-Jordan formula, as stated, is not empirically verifiable. The problem is not just one of limitations to the statistical techniques employed. It is more fundamental. Fields of intellectual activity and participants in those fields, carrying on teaching and research, are intricately bound together. If it makes little sense to conceive of university programs with no students, teachers, or researchers, or of participants with no programs, it makes little sense to postulate that library size is separately influenced by these factors. To account for library size in a causative, explanatory way an entirely different tack must be taken. I offer no solution in this paper but encourage students of library science to develop one, for in doing so they will necessarily evolve a much sharper and more realistic conceptualization of the university library.

A way around the statistical difficulties described above, but one that leads us away from the Clapp-Jordan formula, is to utilize one or the other but not both of the variables F and D. Equations [6] and [7] present the results of regressions which leave out the number of programs.

[6] $$V_1 = -1231 + \underset{(.93)}{} \underset{(3.13)}{.724F} + \underset{(1.05)}{.617S} + \underset{(1.94)}{5.907G} + e$$

$\bar{R}^2 = .22$ $\quad F = 4.30$

[7] $$V_2 = \underset{(2.2)}{-22} + \underset{(5.01)}{.0085F} = \underset{(2.57)}{.0111S} + \underset{(3.02)}{.0671G} + e$$

$\bar{R}^2 = .44$ $\quad F = 10.28$

Neither of these new equations improves the fit of the relationship to account for a greater part of the variability of library size. In that respect there is no improvement. However, the relationship that appears to exist between library size and the size and nature of the university can be given a more satisfactory interpretation. The number of faculty members, taken as a general indicator of the size of the university, exerts a strong influence on library size. The number of volumes rises by 724 for every additional faculty member; the number of current periodical titles by 8½. The coefficient of the F variable is statistically significant in both cases. The additional influence on the number of volumes in the university library appears to relate more to the research and graduate training function of the university rather than the extent to which student enrollment differs from proportionality with faculty size. However, both influences are statistically significant in the determination of the numbers of current periodical titles.

The regression equations, either including or excluding numbers of doctoral fields, give predictions of library size that vary widely from the actual figures for the universities included in the sample. These residual deviations warrant study to see first if they are truly random, as has been assumed in the estimation of the regression equations, or whether they can be associated with any readily recognizable characteristics of universities. One possibly interesting question that comes to mind is whether the unexplained or residual part of library size is related to commonly held views of the quality ranking of graduate schools. That turns out not to be the case. The largest underpredictions, indicating university libraries that are much larger than would be expected on the basis of average practice, are for Harvard, Yale, Illinois, and Duke. Substantial overpredictions are made for Wisconsin, Pittsburgh, and Pennsylvania. The equations that incorporate numbers of doctoral programs produce large overpredictions for Chicago and Johns Hopkins as well. These last, especially, would hardly be regarded as weak graduate schools. The one characteristic which stands out in the residuals from the regression equations is that the equations underpredict for those institutions (e.g., Harvard, Yale, Illinois, Duke) which are renowned for their special attention to libraries. This result gets some corroboration from a positive, although not strong, correlation between residuals from the equation for volumes and that for periodical titles. What this seems to point to is that for some academic institutions the library is more than just a resource for teaching and research but is something of an end in itself. These institutions would presumably justify library collections much larger than would be indicated by their usual research needs, on the grounds that they view as a valid part of their function the preservation of part of civilization's heritage. In that sense, some universities have been prepared to develop national or regional libraries while others have been more content to restrict their ambitions to the needs of teaching and research on their campuses. I would be hesitant to press too far the notion that the positive residuals from the regression equations reflect the extrauniversity or, perhaps more accurately extraresearch, goals of universities. They seem to point in that direction but the random element is too great and the fit of the regression equations too poor to make much of such an argument. No other systematic element appears to be evident in the residuals.

The principal object of this study was to evaluate the Clapp-Jordan formula as a basis for estimating minimum levels of adequacy of academic libraries for graduate studies and research. That is not readily done directly. It was necessary to modify the formula to be able to estimate the relationship between library size and size and diversity of research efforts and graduate training programs by means of regression analysis. It would be inappropriate to attempt to compare directly the coefficients of the regression equations presented in this paper with the weights of the Clapp-Jordan formula. One point that might be made in this regard, though, is that the proportion of students who are in graduate programs—a variable that does not enter into the Clapp-Jordan formula—plays an important role, especially in the reduced version.[13]

A better method of evaluation would be to compare the predictions obtained with the regression equations with those obtained from the Clapp-Jordan formula. This cannot be done with quite as much precision as might be desired since I have not obtained all of the information that would go into the Clapp-Jordan formula. What is missing, though, is the number of undergraduates in honors programs and the number of undergraduate fields of specialization. Where we are dealing with libraries of several million volumes these variables have little influence. For few institutions could they account for more than 100,000 volumes. A more serious matter is the degree of arbitrariness that is involved in counting numbers of fields of doctoral study. I indicated earlier that these could be counted at varying levels of aggregation. My chief concern in preparing data for the regression analysis was to put the listings of fields for all universities on as comparable a basis as possible. It is not precisely clear how Clapp and Jordan handle this problem but they seem to accept university statements about fields at face value without concern for consistency. The weight attached to numbers of doctoral fields is so large that this variable plays the strongest role of any in the formula. It therefore becomes especially important that the variable is adequately defined. In their original presentation, Clapp and Jordan give explicit results for the application of their formula to only three full-fledged universities: Illinois, Michigan, and UCLA. They count twelve more fields for Michigan and nine more for Illinois than I and five fewer for UCLA.[14] Since they do not document precisely how they ascertain the number of fields I have no way of reconciling these counts but I strongly suspect that they have not determined that fields are defined consistently between universities.[15]

Consider first the three schools for which Clapp and Jordan provide explicit estimates. Using my data rather than theirs we get the following formula results:

| | |
|---|---|
| Illinois | 2,163,000 vols. |
| Michigan | 2,226,000 vols. |
| UCLA | 1,723,000 vols. |

My figure for Illinois is well below that shown by Clapp and Jordan and for UCLA it is slightly higher. The main concern here, however, is how these results compare with predictions from the regression equations. For both UCLA and Michigan I get very similar results with either Regression II, incorporating numbers of fields in line with the Clapp-Jordan approach, or Regression I, using only size variables. And in both cases the regression predictions are well above those obtained with the Clapp-Jordan formula. For Illinois, Regression I agrees closely with the predictions from the Clapp-Jordan formula, whereas with Regression II the prediction is 350,000 volumes higher and well above that of the Clapp-Jordan formula.

| | Regression I | Regression II |
|---|---|---|
| Illinois | 2,074,000 | 2,423,000 |
| Michigan | 3,065,000 | 2,995,000 |
| UCLA | 2,378,000 | 2,413,000 |

Similar comparisons can be made for other universities, using the data gathered for my regression analyses. As a general rule the regression equations produce a higher figure for the expected number of volumes than does the Clapp-Jordan formula. This is not surprising since the regression equations endeavor to measure the average relationship, whereas the Clapp-Jordan formula is intended to indicate a minimum standard. Whether the Clapp-Jordan formula indeed points to minimum levels of adequacy is something that cannot be concluded from this analysis. What can be said, however, is that there is nothing to indicate that it produces an overprediction. Applied to those universities that are already heavily engaged in graduate education and research at the doctoral level in a serious way, the Clapp-Jordan formula does not produce results that are patently too high. In the light of suspicions expressed by government officials and budgetary authorities this may be an important conclusion. We must recall, though, that it is a conclusion reached through comparison with the results of regression analysis which was subject to a high degree of variability. It should not, therefore, be unduly emphasized. Still, it may be of some comfort to those who have used the Clapp-Jordan formula in support of claims to build collections of a minimally adequate size that they cannot be accused of excesses. For university research libraries Clapp and Jordan offer a conservative guide. Viewed in that light, the results reported in this paper suggest that as a very rough, quickly computed guide to minimum levels of library size, the Clapp-Jordan formula should remain in the librarian's tool kit. What I hope that this attempt at empirical verification of the formula has shown, however, are the inherent weaknesses of the formula and its sensitivity to definition of fields. More importantly, I hope that it has shown the need for developing predictive formulas from causative explanatory models of the nature of research libraries.

## References


1. Verner W. Clapp and Robert T. Jordan, "Quantitative Criteria for Adequacy of Academic Library Collections," *College & Research Libraries* 26:371–80 (Sept. 1965).
2. Readers unfamiliar with statistical techniques will find an explanation of linear regression analysis in any introductory textbook on statistics with application to business or the social sciences.
3. The particular assumption made here is the common one that they follow a normal distribution.
4. See their "Holdings and Expenditures of U.S. Academic Libraries: An Evaluative Technique," *College & Research Libraries* 27:478–87 (Nov. 1966).
5. Their results for 1952 are not fundamentally different.
6. These were drawn from the list presented by Allen M. Cartter, *An Assessment of Quality in Graduate Education* (Washington, D.C.: American Council on Education, 1966).
7. *Library Statistics of Colleges and Universities 1965–66* (Chicago: American Library Association, 1967).
8. *Students Enrolled for Advanced Degrees, Fall 1966* (Washington: G.P.O., 1968) and *Opening Fall Enrollment in Higher Education, 1966* (Washington: G.P.O., 1967).
9. *American Colleges and Universities* (10th ed., Washington: American Council on Education, 1968).
10. It appears that reasonable uniformity can be obtained with the use of the following classes: Classical Languages, Other Ancient Languages, English, Romance, Germanic including Scandinavian, Slavic, Near and Middle Eastern, and Oriental languages. Since language fields are highly intensive in the use of library resources this categorization may be too aggregative. These fields would probably use a larger body of literature each than many of the natural science fields. However, to be more specific for some universities would require either special information from the universities or assumptions about what subareas are covered.
11. An attempt has been made to approximate full-time equivalent faculty numbers by weighting part-time faculty one-half that of full-time. My first thought was that, in terms of their demands on bibliographic resources for research, part-time faculty should be counted the same as full-time. Indeed, some faculty are regarded as part-time because they have appointments in research institutes. The definition of part-time status seems to vary widely from institution to institution so that one found curious anomalies in student-faculty ratios if part-time faculty were not given a lesser weight.
12. $\bar{R}^2$, is the coefficient of determination, adjusted for the number of degrees of freedom. It is perhaps more directly informative than the commonly used R—the coefficient of correlation—as an indicator of the goodness of fit of the regression equation. I show also the statistic F for the estimated equations in order to provide a test of the hypothesis that the explanatory variables are unrelated in the sense that the regression coefficients are really zero. That hypothesis would be rejected, at a .95 level of confidence, if F exceeds 2.53 for 5 and 30 degrees of freedom.
13. Let us call this for brevity, Regression I. The version which includes numbers of science and nonscience doctoral fields will be called Regression II.
14. Clapp and Jordan, "Quantitative Criteria."
15. I indicated previously that I considered the argument that a doctoral program is one that a university sees fit to declare, but this approach must be unsatisfactory since it leads to results that are difficult to interpret. Administrative practices and views about the appropriate scope of disciplines vary widely, especially in the organization of doctoral programs.