

Reconciling Pragmatism, Equity, and Need in the Formula Allocation of Book and Serial Funds

Charles B. Lowry

The literature of formula allocation has several central themes which the matrix formula described here addresses. Most important, formulas have almost universally failed to provide a mechanism to distribute funds for serials as well as books. The matrix formula allocates funds for monographs and serials based on disciplinary needs and publishing patterns. It also provides a method for determining the variables which best represent institutional goals, normalizing them and explicitly determining the percent of funds allocated by the individual variable. These features are great advantages in dealing fairly with the difficult problems of allocating scarce resources. Finally, the article discusses the appropriate limits formulas may have as allocation tools.



about ten years ago a colleague responded to one of my enthusiastic discussions of what I will call the matrix formula with a comment that "all formulas are to a certain extent derivative." It is fair to ask why anyone should revisit the topic of formula allocation of acquisition funds, given the extensive discussion of it in the literature. Keeping my colleague's response in mind, I believe that gains are to be made in the allocation of scarce resources if librarians build on earlier work and find new ways to reuse what has been accomplished. This paper does not review the literature of formula allocation, but rather points to examples of earlier work that inform the matrix formula or to problems that the matrix formula arguably resolves. For readers who wish to explore past research, the review essay by Mary Sellen, published in *Collection Management*, is a fair place to start for discussion of what has been done. Both Donna Packer's and Fabian A.

Ehikhamenor's articles also provide a window on the literature of the subject.¹

John Budd and Kay Adams's recent study, which involved a survey of 834 academic libraries (43% return rate), indicates that libraries' use of allocation formulas has declined. In the 1940s about 73% of college libraries indicated a use of formulas whereas today only 41% "report using some sort of formulation for the allocation of materials resources."² Nonetheless, formula allocations are clearly going to remain a part of the library scene, and I believe they are a preferable alternative to leaving things to chance, incremental budgeting, or historical precedent, so long as librarians understand the specific limitations that characterize this budgeting mechanism. This is not to say that historical precedent or incremental budgeting fail on all counts as funding mechanisms, but that they may particularly discriminate against disciplines with rapidly changing literatures or newly established curricula and

research emphases. For instance, recently there has been a call among psychologists to move toward collection use, publication cost, and similar empirical data as part of the criteria for distributing acquisition funding, exactly as formulas have often done. Some psychologists assert that research libraries defend their fund allocations to various disciplines by comparing one discipline to another. This process effectively bars change and places psychology at a disadvantage.³

In collection development libraries have numerous objectives that may not be adequately addressed by a formula approach like the one described here. In spite of varied objectives, a formula can be useful for allocating part of the budget. For instance, at the University of Texas at Arlington (UTA) the funds allocated to reference collections, special collections, media, foreign languages, documents, and "enrichment" (largely retrospective purchases by bibliographers) all lie outside the formula. Similarly, a large research library collecting comprehensively in specified subject areas may allocate to these areas outside of the context of a formula. Yet a small or medium-sized academic library may allocate effectively a large part of its budget through the formula process. At UTA in fiscal year 1990-91, 74% of the \$1,750,000 acquisitions budget was distributed through the matrix formula.

On the one hand, the adoption of a new formula is often a by-product of some budgetary crisis that compels a change in materials funding. On the other hand, the numerous obstacles to the introduction of a formula or to changing an existing one are well known. Institutional interest group politics, the open decision-making process practiced in the academy, and simple inertia are all forces that have a significant role influencing the development of formulas. Jasper Schad in the late seventies and Donna Packer more recently both addressed at length how these forces may interact.⁴ Those who benefit from the status quo will not want to investigate alternative budgeting

mechanisms. However, those who are disadvantaged by existing policies may assume that nothing will change and may not make any effort to abet the adoption of a new method. A key problem, both inside and outside the library, is the confusion that inevitably occurs in distinguishing between means and ends. Everyone wants to begin at the end—that is, to know what the formula will give them in real dollar terms. The natural focus of a particularist's interest is selfish. The general good, that is, the general ends of equitable distribution of limited resources, the balanced growth of the collection, and the meeting of *bona fide* instructional or research goals of the institution can be lost in this process.

Several years ago Jasper Schad addressed what he termed "justice principles" for fairness in allocation of book funds. By inference these principles may also be applied to serials. "The principle of need specifies that funds be distributed according to the particular requirements of each discipline . . . ; the principle of contributions stipulates that funds be apportioned on the basis of the degree to which each department serves the institutional mission . . . ; [and] the principle of equality affords each recipient an equal share, regardless of differences in need or output."⁵ The matrix formula can be used to reconcile these competing principles.

Deciding on a formula may itself be a daunting undertaking. Rare is the faculty member, or for that matter the librarian, who is willing to digest the mountain of published literature necessary to developing a comprehensive plan that satisfies the general institutional interests. Often the formulas discussed in the literature are akin to Rube Goldberg contraptions, fascinating to watch, but overly elaborate means to ends. The rule should be to avoid spurious sophistication in the development of formulas. William E. McGrath, in his often quoted article, stated that "formulas must make sense. They must be based on sound theory backed up by empirical data. The difficulty is that no such theory has yet been presented."⁶ I generally find

more in McGrath's article to agree with than not, but to call for a theory of allocation formulas seems to me to miss the point. Gary M. Shirk was even more explicit in his call for "A defensible theoretical framework that relates objective variables to the collection's performance in a meaningful way."⁷ Few such attempts at theory building have been made. A notable exception is James A. Yunker and Carol G. Covey, who attempted to "build on prior work by means of operationalizing the economic approach. This requires the specification of an explicit hypothesis concerning the affect [sic] of library operations on social welfare which utilizes measurable dimensions of these operations."⁸ Yunker and Covey use a theory of social utility as a basis for building a library formula.

However, I believe that the theory called for will never emerge, because allocation formulas are a mathematical means for distributing fiscal resources. The belief that there is a theory waiting to be found is at best an a priori assumption or at worst a misuse of the generally accepted notion of social science theory, which is usually an explanatory theoretical hypothesis such as the "theory of the leisure class" or the "quantity theory of money." In this sense, theory is no more than a generalized explanation of observable social phenomena. Quantitative empirical verifiability may not even be an objective in many cases of social science theory.

In the strictest sense, allocation formulas are—to use T. S. Kuhn's concept—paradigms that explain the empirical basis for distributing library material funds. As such, they become tools for budget decision making, not hypotheses for theory building. Kuhn, a physicist and historian of science, used the concept of paradigm to challenge the predominant conception of the philosophy of science, the hypothetico-deductive view of Karl Popper, and to forward the notion of a "normal science" as it is really practiced by scientists. But the concept of paradigm was quickly embraced in the social sciences and is used here in an applied sense. Philosophically speaking, a formula qua paradigm is an "artifact which can be used as a puzzle-solving device; not a metaphysical world view." "The paradigm is something which can function when the theory is not there."⁹

Formulas generally take care of basic current needs, not retrospective collection development. They are difficult to apply to some disciplines, such as foreign languages, and to interdisciplinary departments; they are frequently difficult to understand without considerable study. Two examples from Budd's recent article are illustrative of the problem (figure 1 and figure 2).¹⁰ These formulas may be quite adequate. But they require considerable study if anyone is to comprehend the interrelationships of the variables and their net effect on allocations. This inaccessibility is an obstacle

$$\text{FORMULA} = [.157446 + .0024(B) + .061(C) + .1979(D) + .0135(E) + .000005(F) + .0003(G) + .0445(H)] / 1.7 = \text{department percentage}$$

B = number of undergraduate majors

C = number of graduate majors

D = number of FTE faculty in the department

E = number of 5-hour courses taught by the department

F = number of volumes in the library in the specific subject area

G = number of volumes circulated in the group identified in F

H = average price as reported in *Publishers Weekly*

FIGURE 1

Figures 1 and 2 are illustrative of the diversity of formulas currently in use and the obstacle their complexity presents to ready understanding.

Unit Cost

$$u = B \left[\text{sch}^1 \left(\frac{10}{\text{SCH}^1} \right) + \text{sch}^2 \left(\frac{20}{\text{SCH}^2} \right) + \text{sch}^3 \left(\frac{20}{\text{SCH}^3} \right) + c \left(\frac{30}{C} \right) + \text{co} \left(\frac{10}{\text{CO}} \right) + f \left(\frac{10}{F} \right) \right]$$

Final calculation

$$D = \frac{A}{u} \times u + \$1,000$$

A = Total nonserial allocation

B = Bowker average price per volume

* c = Circulation

* co = Course offerings

D = Departmental allocation

* f = Faculty

* sch = Student credit hours

¹ lower division

² upper division

³ graduate

* u = Unit cost.

FIGURE 2

to determining precisely how the formulas relate to the collection development goals of the library.

INSTITUTIONAL POLITICS AND RESOURCE ALLOCATION

Whether faculty are deeply involved in selection (as they are at UTA) or selection is totally in the hands of librarian bibliographers, decisions must be made about the allocation of funds to disciplines. This process may be the result of the piecemeal accretion of numerous decisions about individual titles or the prior allocation of a specific amount of acquisition funds for each discipline. Formulas for allocating library materials budgets have developed for two primary reasons. The first, the need to allocate limited resources on a basis of equity and need, is commonly understood and broadly discussed in the literature. The second reason for applying a formula is little discussed in the library literature. Formulas are extremely useful in solving the political dilemmas of

unequal resource allocation. No library has the unlimited resources needed to purchase all the available relevant materials. It follows that the politics of allocating materials funds is critical to the process of dividing resources, whether funds are allocated to individual departments or to subject bibliographers for the purpose of collection development. In either case, the dynamics of local politics play an important, if not dominant, role in how funds are allocated.

Over time, patterns of influence, personal connection, organization, and historical anomalies combine to influence allocation of resources in ways which may or may not meet the institutional mission of the library. In combating the vagaries of these influences, libraries have traditionally turned to formulas that emphasize the use of "rational" or "empirical" variables as a basis for resource allocation. Variables include factors of institutional program size and productivity, such as number of faculty and students or credit hour production

and extra-institutional measures like recent price experience in the book and journal publishing universe. Compelling arguments can be made for the legitimacy of such variables as part of an allocation formula, but emphasis on them has tended to reverse the order in which we think about formulas. Discussions in the literature rarely proceed from the development of a methodological model or explanatory paradigm—that is, the formula—to the process of deciding what variables should be used and where we should get data to support them. Often this inversion is because we have *not* explicitly and overtly accepted the fact that allocation is inherently political in nature with a result that variables and supporting data which have the highest likelihood of acceptance by the institutional units or individuals with a stake in the process are accepted. This process means that the methodology of the formula itself and the relationship of variables to institutional goals are ignored or given insufficient attention.

The alternative course suggested here is to accept the political component of the budgetary allocation process and, accordingly, to identify a model that is acceptable to those factions involved in the budgetary process. The faculty advisory committee, the library's collection development group, or whoever else is involved in the allocation of funds will be better served by a process requiring an explicit articulation of the factors that will be used in determining distribution of fiscal resources. The matrix formula is a context for considering the philosophical case to be made for any variable and at the same time knowing with precision who gains or loses by its inclusion.

STRUCTURE OF THE MATRIX FORMULA

The matrix formula is largely the product of cooperative efforts between the libraries and faculty library committees of four institutions—University of North Carolina, Charlotte; Elon College in North Carolina; University of South Alabama; and the University of Texas at Arlington. These efforts began over fifteen years

ago. The formula has been largely evolutionary in its development, and its features have been positively influenced by the continued scrutiny and discussion of librarians and faculty who had a large stake in making it work effectively. The basic framework of this paradigmatic formula is transportable to many institutions, but within the formula each institution devises variables that fit local needs.

The term *matrix* was chosen because it describes the grid showing each of the variables and their distribution among the various departments. An early version that allocated only book funds to departments at Elon College is shown in table 1. The variables are fairly obvious as is the weighted percent share of each variable by the departments. One of the consistent mathematical problems with formulas is that diverse variables are thrown together willy-nilly without giving consideration to the fact that they are apples and oranges.¹¹ The simple mathematical approach to normalization is to reduce each variable to a percentage of the total for the raw data. For instance, the faculty full time equivalent (FTE) index and the freshman sophomore credit hours (table 2) for thirty-two different departments at UTA (1987–88 academic year) are converted to straight percentages and thus become comparable. This mathematical operation allows any variable to be compared arithmetically to any other variable.

The representation of a variable as a normalized percentage is similar to the use of the raw-data variable in one very important respect. They are both mathematically linear in their effect. The brief discussions in our literature do not do justice to the topic of tempering variables through the use of logarithms.¹² The basic reason for moving from a linear to a logarithmic approach relates to the aphorism that "what counts easiest should not count most."¹³ The examples of table 2 are illustrative of the problem. Several departments have extraordinarily high productivity levels in freshmen and sophomore credit hours. For instance, the Departments of English, History, and

TABLE 1
ELON COLLEGE BOOK ALLOCATION FORMULA—SPRING 1979

Department Variable	Business and Economics	Education and Psychology	Fine Arts	Human Services	Language /Literature	Philosophy	Physical Education	Religion	Social Sciences	Biology	Total
% books published x5	41.6	56.55	14.45	5.15	99.45	15.30	35.85	31.70	169.50	30.45	500
% average book cost x4	41.2	36.44	41.60	29.52	30.28	44.92	40.40	29.76	38.20	67.68	400
% FTE faculty x3	41.07	24.36	37.11	8.67	73.89	7.86	22.38	15.72	47.97	20.97	300
% number students at 100-400 course levels x2	38.48	21.04	12.76	4.78	40.98	4.36	27.38	11.44	30.32	8.46	200
% graduating seniors x1	38.79	8.59	3.32	10.80	3.32	.27	9.70	1.66	13.58	9.97	100
Total weighted value	201.14	146.98	109.24	58.92	247.92	72.71	135.71	90.28	299.57	137.53	1500
% book allocation	13.41	9.8	7.29	3.93	16.53	4.84	9.04	6.02	19.97	9.17	100

Mathematics range from 21,000 to more than 35,000 credit hours for the year in question. A strong case can be made that as the number of credit hours generated increases, particularly in large classes of service courses with many sections, there is a diminishing need to provide book funds to support credit-hour production. Put another way, in the allocation of acquisition funds, the credit hours produced by the first student to enroll in a large lecture class should count far more than that for the 251st student. Figure 3 shows a simple linear and logarithmic plot of the same increasing values of 1-10. The logarithmic plot dramatically dampens the effect of the increase in the size of the number. Figure 4 illustrates what happens when logs are applied to the raw variable of credit-hour production. Similarly, table 3 shows the conversion of raw credit-hour production to logged data and finally to a logged percent. The diminishing effects of this conversion on the larger CHP numbers are evident when comparing the logged to the linear percent.¹⁴

Weighting is the other mathematical transformation of variables that should be discussed more broadly in the literature. "In much of the literature there is an implicit acceptance of an equal weight for each of the variables in the formula."¹⁵ Nevertheless, weighting is an acceptable technique for modifying the base value of variables so that they are unequal. Decisions about weights determine the extent to which Schad's principle of equality will be applied. The matrix formula under discussion here provides an explicit mechanism for weighting each variable, and the weighting is a distinctly understandable measure of the percentage of the funds allocated by that variable. Table 4 illustrates the weighting and the log value given to the variables presently in use at the University of Texas at Arlington Libraries. These weights and logs emphasize the point made earlier, that a formula should reflect the collection development goals of the institution. The explicit understanding of the effects of any one variable is made clear in a matrix formula. For in-

TABLE 2
ILLUSTRATION OF LINEAR DATA NORMALIZATION

	Faculty FTE Index	Faculty FTE	Freshman-Sophomore	Freshman-Sophomore
	Raw Data	Index Linear Percentage Data	Credit Hours Raw Data	Credit Hours Linear Percentage Data
Architecture	43.24	4	5,767	2
Art	23.34	2	2,553	1
Accounting	38.19	4	9,855	4
Economics	28.80	3	9,948	4
Finance	29.76	3	207	0
Management	29.96	3	6,075	2
Marketing	19.08	2	0	0
Systems analysis	38.42	4	5,493	2
Aerospace engineering	22.24	2	3,461	1
Civil engineering	28.87	3	3,484	1
Computer science	40.15	4	6,104	2
Electrical engineering	48.78	5	3,830	2
Industrial engineering	14.69	2	0	0
Mechanical engineering	28.08	3	729	0
Education	16.55	2	650	0
Communications	30.10	3	11,165	5
English	48.91	5	26,883	11
History	33.23	3	20,265	8
Music	28.54	3	5,820	2
Philosophy	8.23	1	3,351	1
Physical education	21.46	2	11,684	5
Political science	28.54	3	18,840	8
Sociology	27.02	3	3,729	2
Nursing	52.89	5	0	0
Biology	36.27	4	15,350	6
Chemistry	35.04	4	10,196	4
Geology	17.56	2	2,344	1
Mathematics	55.00	6	34,435	14
Physics	32.43	3	15,045	6
Psychology	27.43	3	6,014	2
Social work	33.46	3	1,038	0
Urban affairs	9.81	1	0	0
Totals	976.07	100%	244,315	100%

TABLE 3
ILLUSTRATION OF LOGARITHMIC DATA NORMALIZATION

	Freshman-Sophomore Credit Hours Raw Data	Freshman-Sophomore Credit Hours Logged Data	Freshman-Sophomore Credit Hours Logged Percentage	Freshman-Sophomore Credit Hours Linear Percentage Data
Architecture	5,767	0.53	3	2
Art	2,553	0.31	2	1
Accounting	9,855	0.70	4	4
Economics	9,948	0.71	5	4
Finance	207	0.04	0	0
Management	6,075	0.54	3	2
Marketing	0	0.00	0	0
Systems analysis	5,493	0.11	3	2
Aerospace engineering	3,461	0.38	2	1
Civil engineering	3,484	0.38	2	1
Computer science	6,104	0.54	3	2
Electrical engineering	3,830	0.41	3	2
Industrial engineering	0	0.00	0	0
Mechanical engineering	729	0.11	1	0
Education	650	0.10	1	0
Communications	11,165	0.75	5	5
English	26,883	1.08	7	11
History	20,265	0.97	6	8
Music	5,820	0.53	3	2
Philosophy	3,351	0.38	2	1
Physical education	11,684	0.76	5	5
Political science	18,840	0.94	6	8
Sociology	3,729	0.40	3	2
Nursing	0	0.00	0	0
Biology	15,350	0.86	6	6
Chemistry	10,196	0.71	5	4
Geology	2,344	0.29	2	1
Mathematics	34,435	1.18	8	14
Physics	15,045	0.85	5	6
Psychology	6,014	0.54	3	2
Social work	1,038	0.15	1	0
Urban affairs	0	0.00	0	0
Totals	244,315	15.57	100%	100%

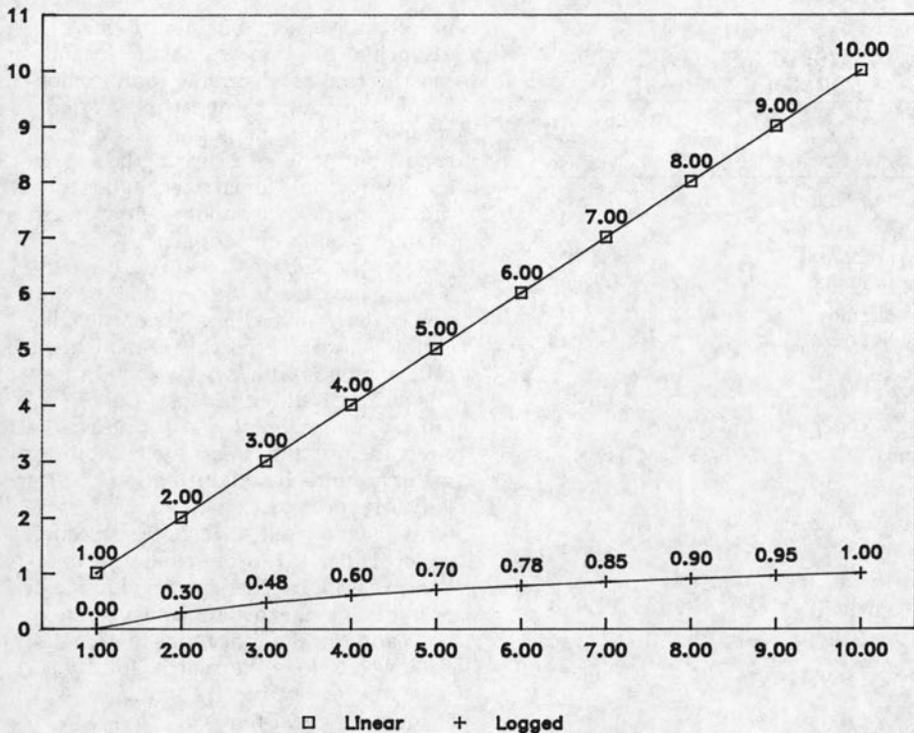


FIGURE 3. Linear versus Log Base 10 Plots

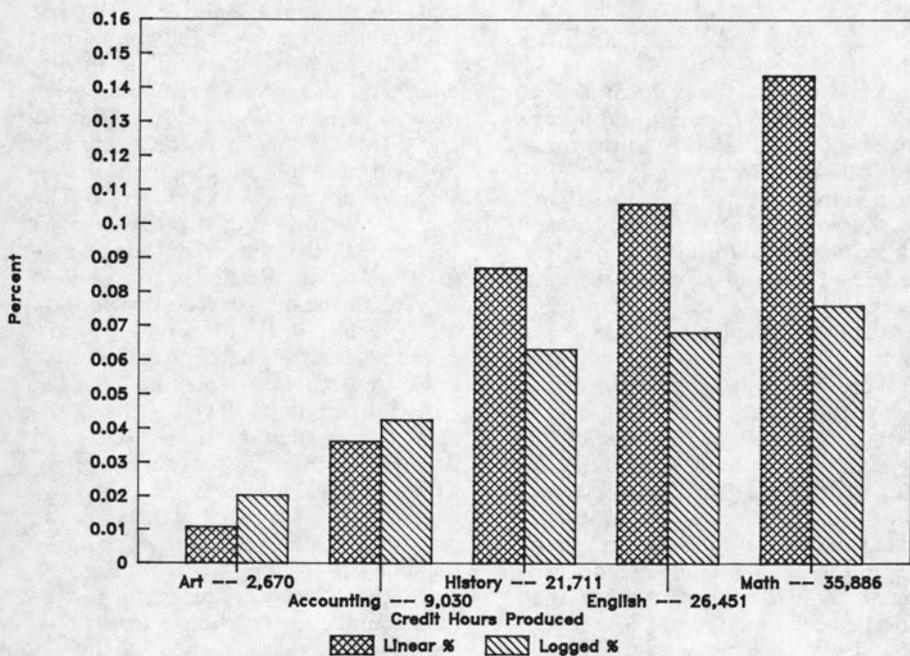


FIGURE 4. Linear versus Logged Percents of CHP

TABLE 4
UTA LIBRARIES ALLOCATION
FORMULA VARIABLES

Variable Name	1988-89 Weight %	Logged or Linear Variable
Faculty FTE Index	10	Linear
Freshman- sophomore credit hours	1	Logged
Junior-senior credit hours	1	Logged
Master's credit hours	3	Linear
Doctoral credit hours	12	Linear
Library Use Index	5	Logged
Graduate degrees index	5	Linear
Undergraduate degrees index	3	Logged
Book cost index	21	Linear
Serials cost index	21	Linear
Publications index	15	Linear
External grants index	3	Linear
Total	100	

stance, the emphasis on doctoral and postdoctoral work, research, and the acquisition of grants at UTA sums for a total of 30% of the formula, that is, 30% of the dollars are allocated by these variables. Moreover, all of these variables are linear in the present formula so that they have a direct effect on the dollar outcome proportional to the individual departmental share of the variable.

Another unique feature of the matrix formula is the Book/Serial Dependency Index. Within the formula it has several important functions that help tailor the outcomes of the formula to the idiosyncracies of different disciplines, thus adhering to Schad's "principle of need." The size and cost of the book and serial literature of each discipline frequently dominate the discussion of variables that should be used in a formula. The use of book and serial pricing as a variable will

be discussed later, but here it serves a distinctive function within the formula. Most formulas distribute only monograph allocations to departments. A few actually allocate monograph and serial funds, but none give direct guidance in the individual department allocation proportions which should be expended on monographs and serials.¹⁶

Recently, Robin B. Devin and Martha Kellogg have made the case that the citation patterns in the literature of each discipline should be used to establish the ratio of monograph/serial expenditures. Devin and Kellogg fall into the logical fallacy *pro hoc, propter hoc*—putting the effect before the cause.¹⁷ Distributing material funds based on citation studies will not produce those same citation patterns. It may well disturb the collection patterns that led to the citation proportions. Distributing departmental funds in proportion to the total monetary value of book and serial publications in each discipline—as does the matrix formula—seems more empirically sound, as well as logically consistent.

The effects of the Book/Serial Dependency Index are demonstrated in table 5 (see pages 132 and 133). The first two columns represent the annual cost for both formats by discipline during one year. This is directly computed as the percentage dependency index (columns 3 and 4), which then is used to compute the distribution of the departmental allocation between books and serials. The use of this proportionality is based on the rationale that the cost of the universe of scholarly publications more or less represents the appropriate distribution of expenditures of local funds. Moreover, it is arguable that the actual publication of monographs and serials in a discipline reflects the structure of the discipline's communication process and results in the characteristic citation pattern that the discipline evinces.

Vital to this approach is the accurate, dependable, and consistent collection of the data supporting it. This very important data-collection process will be discussed in the next section. Surprisingly, the proportions established by the de-

pendency index have gained easy acceptance by faculty and the library advisory committees at both institutions—the University of South Alabama and the University of Texas at Arlington—where I have applied them over the last ten years. The total UTA dependency index is heavily weighted toward serial expenditures. This weighting does not necessarily result in an equivalent distribution of funds due to the use of other weighted variables. Not surprisingly, science disciplines have tended to expend their full serial allocations under the formula. Although scientists have accepted the limits the formula dictates, they have still expressed a strong desire to add more titles to the subscription list when other disciplines have not. At present, departments with serial obligations that exhaust their allocation cannot add new serial titles without canceling title subscriptions of equal value. Incidentally, this practice requires that the UTA Library assign every title in the serial collection to either a department, a college, or the library's own general serial fund.

CHOOSING VARIABLES— SOME COMMENTS

The choice of variables and the collection of the data to support them are the dominant themes of the literature on formula allocations. Indeed, the implication often is that the variables are equivalent to the formula. The contention here is that the adoption of an explicit mathematical model or formula paradigm is the first and perhaps the most important step in developing an allocation formula, and that the choice of variables which follows should be dictated by the institutional objectives. Variables should be chosen in light of local conditions and institutional goals. The selection of variables is important, and plenty of discussion in the literature attests to this fact.¹⁸

One of the key issues in variable selection is the dichotomy among variables that may be chosen. Sweetman and Wiedemann refer to this as the supply side and demand side.¹⁹ Put simply, this is the difference between internal variables generated by programs within the insti-

tution and external variables such as the price of library materials or the citation characteristics of the individual disciplines.²⁰ A large variety of variables are in use in library allocation formulas, but a handful have clearly become dominant, including "number of students or number of student credit hours, cost of materials, number of faculty, circulation by department or subject area, number of courses offered by a department, and the number of students majoring in a department or subject area."²¹ Most of these variables involve Schad's principle of contribution, and the matrix formula allows their dollar effect to be explicitly specified.

The matrix formula gives more or less complete flexibility in developing variables to fit the local situation.

At present, twelve variables are being used in the matrix formula at the University of Texas at Arlington (table 4). There is a point at which selecting new variables to "enhance" the matrix formula has a diminished result. As the number of variables increases, the weighted percent share must be reduced. New variables are usually suggested to correct some perceived oversight, but twelve is probably the practical limit. Substituting new variables for old ones is probably better when that limit is reached. It bears repeating that the matrix formula gives more or less complete flexibility in developing variables to fit the local situation. To illustrate, I have often joked that there could be a variable for blue-eyed full professors in each department if that were perceived as an appropriate criterion for distribution of materials funds.

Two variables merit brief discussion: the Library-Use Index, an internal variable, and the Book/Serials Cost Index, an external variable. The literature discusses Library-Use Variables a great deal and explores the numerous inherent problems. I think that we have addressed this variable at UTA in a consistent and logical

TABLE 5
UTA BOOK/SERIAL DEPENDENCY INDEX

	Book Cost Index (BNA Universe)	Serials Cost Index (FAXON Universe)	Book Dependency Index(%)	Serial Dependency Index(%)	Current FY Book Budget Allocation	Current FY Serial Budget Allocation
Architecture	\$ 26,406.00	\$ 17,794.42	59.74	40.26	\$ 14,682.12	\$ 9,893.96
Art	58,316.00	46,538.03	55.62	44.38	11,035.68	8,806.83
Accounting	6,514.00	14,113.60	31.58	68.42	7,629.88	16,531.32
Economics	49,615.00	233,471.66	17.53	82.47	8,061.22	37,933.42
Finance	5,358.00	141,522.63	3.65	96.35	1,688.23	44,591.86
Management	58,522.00	173,433.49	25.23	74.77	12,674.59	37,561.91
Marketing	5,021.00	41,742.09	10.74	89.26	2,662.40	22,133.85
Systems analysis	31,776.00	44,715.50	41.54	58.46	16,749.81	23,570.50
Aerospace engineering	23,845.00	18,185.63	56.73	43.27	13,084.59	9,979.09
Civil engineering	15,262.00	145,985.03	9.46	90.54	4,192.43	40,101.70
Computer science	9,554.00	61,508.98	13.44	86.56	5,164.77	33,250.97
Electrical engineering	35,955.00	125,470.14	22.27	77.73	14,369.10	50,143.05
Industrial engineering	27,573.00	24,166.88	53.29	46.71	13,673.25	11,984.18
Mechanical engineering	20,716.00	172,665.29	10.71	89.29	5,743.77	47,873.62

TABLE 5 (continued)
UTA BOOK/SERIAL DEPENDENCY INDEX

Education	30,865.00	58,285.52	34.62	65.38	7,597.26	14,346.00
Communications	20,008.00	29,002.40	40.81	59.19	6,227.24	9,032.86
English	115,726.00	11,892.98	90.68	9.32	34,262.07	3,521.06
History	98,392.00	93,256.68	51.34	48.66	23,274.33	22,059.58
Music	18,014.00	17,995.70	50.03	49.97	5,165.54	5,160.30
Philosophy	57,722.00	34,906.65	62.32	37.68	10,804.07	6,582.00
Physical education	16,209.00	18,143.86	47.18	52.82	6,685.52	7,483.57
Political science	71,470.00	135,198.12	34.58	65.42	13,470.34	25,481.52
Sociology	68,103.00	54,159.99	55.70	44.30	19,043.05	15,144.29
Nursing	36,116.00	133,434.43	21.30	78.70	11,489.30	42,448.45
Biology	118,186.00	510,387.60	18.80	81.20	17,458.30	75,393.85
Chemistry	41,593.00	409,450.84	9.22	90.78	8,119.39	79,929.05
Geology	24,123.00	132,030.96	15.45	84.55	4,781.73	26,171.57
Mathematics	40,741.00	130,011.60	23.86	76.14	15,742.57	50,237.26
Physics	76,631.00	283,301.98	21.29	78.71	13,775.46	50,927.37
Psychology	35,351.00	104,420.50	25.29	74.71	9,616.99	28,406.85
Social work	19,666.00	40,143.76	32.88	67.12	22,639.76	46,214.03
Urban affairs	19,187.00	30,284.11	38.78	61.22	10,261.81	16,196.89
Totals	\$1,282,536.00	\$3,487,641.05	26.89	73.11	\$371,906.56	\$919,093.44

fashion that resolves the empirical problems. First, we count all instances in which library materials are handled, whether during external circulation or reshelving inside the library, as a single unit of use representing a real demand or need. This practice means that a patron's browsing of a book or journal inside the library is equal to checking an item out. We make no attempt to determine who used library materials but rather assign the use by classification number to a specific department. We assume that the department that selected the materials should get credit for all use, whatever the discipline of the patron using the material. To use a simple example, if a student in the department of economics checks out John Kenneth Galbraith's *The Great Crash*, then the history department gets credit because it ordered the book. The collection of these statistics has gradually become easier as the library collection has been retrospectively converted. Once we complete the bar-coding project for our serials, we will be able to capture the internal browsing statistics and the external circulation through NOTIS as part of the reshelving effort and extract the data by generating reports from the circulation subsystem.

Collecting data on the cost of the literature is vital to the matrix formula in order to generate the Book/Serial Dependency Index.

I cannot state too strongly that collecting data on the cost of the literature is vital to the matrix formula in order to generate the Book/Serial Dependency Index. At UTA it is equally important as one of the key variables in our application of the formula since it is weighted to produce 42% of the allocation of funds. If all department programs were equal—i.e., the same enrollment, credit-hour production, and degree level—then this would be the only variable needed. Since the dependency indexes were first incorporated into the formula ten years ago,

determining a source for prices has been a continuing problem, with many authorities consulted including *Publisher's Weekly*, the *Bowker Annual*, periodical tabulations in *Library Journal*, and laborious hand calculations from *Ulrich's*.²² By 1978, the approval plan statistics of Blackwell North America were being published regularly. These represent as close to a comprehensive list of the annual output of academic library books in the English language as can be expected. The subject section of the BNA statistics is used as a basis for dividing this pricing information among the disciplines. Finding a similar statistical universe in published form for serial literature is next to impossible. At the University of South Alabama we gathered statistics by assigning all appropriate titles listed in *Ulrich's* to the various disciplines. At UTA we extracted the data from the entire FAXON database by assigning all titles to appropriate departments and then compiling their dollar value. This data collection has been accomplished manually in the past, but beginning in fiscal year 1989-90, with the help of UTA's Academic Computing Center, we developed a computer method for this purpose using data tapes provided by FAXON.²³ Computerization allows us, with far less effort, to assemble quickly the matrix formula's Serial Dependency Index from year to year and, incidentally, to look at increases in pricing on a title-by-title basis. On the other hand, the Book/Serial Dependency Index is vital both as a mechanism within the formula and as a variable for allocating funds. On the other hand, it is equally important to remember that the data used in the index at present primarily include costs for English-language books of North American and British origin and English-language serials of North American and West European origin.

A survey of the literature on variables shows one important lacuna, which is the need for allocation formulas to take into account somehow the collection-development policy of the individual institution. For example, a variable based on the Resources and Technical Services Division (now ALCTS) definitions of col-

lection levels (teaching, advance research, etc.) might make it possible to continue building in defined areas of excellence. Similarly, analysis of weak areas in the collection could be built into a variable that would provide extra funding for a period of years to ensure that weak areas could "catch up." These are problems we are just beginning to tackle as the formula matures at UTA. Collection of variable data requires consistency from year to year, a public and open process, and accountability for data represented in the formula. New variables may be difficult to introduce because of the commitment that develops around the formula once established. We have used a strategy of assigning variables very small weights so that they gain acceptability, for instance, in the cases of the publication/research and grants variables at UTA.²⁴ It was argued in the Advisory Committee that research and publication take several years to develop. Accordingly, the variable data capture a five-year moving window of faculty productivity in refereed publications. This information is extracted from UTA's annually published list of "Faculty Publications and Creative Activities."

CONCLUSIONS

We expect and encourage the use of the matrix formula outside UTA. On request we have distributed copies of the LOTUS 1-2-3 formula programming to over a dozen other institutions. We have prepared a template program to facilitate use of the matrix formula by other institutions. We ask only that other libraries using the matrix formula give proper public credit and not distribute the formula further.

The structure of the matrix formula has a number of important features. In the first place, it allows total flexibility in the choice of variables. Any type and any number of variables can be used because they may be normalized on a linear or logarithmic base. In addition, each variable may be assigned a percentage weight that makes the dollar outcomes resulting from its use explicit to anyone who examines the formula, thus removing one

of the complaints about the complexity of formulas. Additionally, administrators, faculty, and librarians can shift their energies from the often politically charged discussion of variables to the more fruitful questions of the institutional priorities involved. Variables may be changed from time to time in order to adapt the formula to changing needs in the institution.

Moreover, the formula helps to deal head-on with the statistical problem of multicollinearity and the resulting hunt for the surrogate variable, the "Holy Grail" of variables. Multicollinearity refers to the fact that some variables are basically measures of the different dimensions of the same thing. A glance back at table 4 is illustrative. In some sense, the variables FTE index, credit hours, the graduate degrees index, and the undergraduate degrees index all are representative of the size of specific disciplines at UTA and involve Schad's contribution principle. The literature devotes a great deal of statistical sound and fury to this issue. Within the matrix formula, the question is rendered moot by the fact that, while these variables represent similar measures of size, their specific effect or weight in the allocation of funds is clearly known. We determine the importance of each variable in our institutional setting by its weight and by whether the variable is logged or linear, allowing the question of collinearity to be put to the side. The total weight of the size of program, or contribution, variables in this case is 35%, and we do not concern ourselves about which is the best surrogate or proxy variable to use.

Finally, the formula allocates to each department book and serial funds. What is more, the formula gives a specific quantitative guide to the appropriate distribution of funds between serial and monographic expenditures in each discipline. Thus, the matrix formula is a powerful tool in dealing with the inexorable demand that serial expenditures place on the library budget because it gives a benchmark against which a department serial budget may be limited. At present, only eight of our departments are expending all of their serial

allocation, and we have tended to be conservative in allowing departments to expand their serial expenditures so that others do not reach their limit. Unexpended serial funds may be used by the department for other material expendi-

tures. However, in the end the matrix formula cannot overcome the central problem of budgeting—insufficient funds. Any formula may satisfy acquisition needs if sufficient funds are put into it. No formula will overcome inadequate funding.

REFERENCES AND NOTES

1. Mary Sellen, "Book Budget Formula Allocations: A Review Essay," *Collection Management* 9:13-24 (Winter 1987); Donna Packer, "Acquisitions Allocations: Equity, Politics, and Formulas," *Journal of Academic Librarianship* 14:276-86 (Nov. 1988); and Fabian A. Ehikhamenor, "A Formula for Allocating Book Funds. The Search for Simplicity and Flexibility," *Libri* 33:148-61 (Copenhagen: Munksgaard, 1983).
2. John M. Budd and Kay Adams, "Allocation Formulas in Practice," *Library Acquisitions: Practice & Theory* 13:381-84 (1989).
3. "Are the Books Stacked against Psychology?" *APS Observer* 4:1, 10, 12-13 (May 1991).
4. Jasper G. Schad, "Allocating Materials Budgets in Institutions of Higher Education," *Journal of Academic Librarianship* 3:328-32 (Jan. 1978); and Packer, "Acquisition, Allocations," p.279-82.
5. Jasper Schad, "Fairness in Book Fund Allocation," *College & Research Libraries* 48:481 (Nov. 1987).
6. William E. McGrath, "A Pragmatic Book Allocation Formula for Academic and Public Libraries with a Test for Its Effectiveness," *Library Resources & Technical Services* 19:356 (Fall 1975).
7. Gary M. Shirk, "Allocation Formulas for Budgeting Library Materials: Science or Procedure?" *Collection Management* 6:46 (Fall/Winter 1984).
8. James A. Yunker and Carol G. Covey, "An Optimizing Approach to the Problem of Interdepartmental Allocation of the Library Materials Budget," *Library Acquisitions: Practice and Theory* 4:203 (1980).
9. Margaret Masterman, "The Nature of a Paradigm," in *Criticism and the Growth of Knowledge, Proceedings of the International Colloquium in the Philosophy of Science, London, 1965*, V. 4, ed. Imre Lakatos and Alan Musgrave (Cambridge: Cambridge Univ. Pr., 1970) p.59-89. This essay is a very accessible discussion of Kuhn's ideas, but focuses primarily on their implications for science. For a discussion of the use of paradigm in a social science, see David Hackett Fischer, *Historians' Fallacies, Toward a Logic of Historical Thought* (New York: Harper, 1970), p.161-63.
10. Budd and Adams, "Allocation Formulas," p.388.
11. See, for instance, Packer, "Acquisitions Allocations," p.283. The values for the western Washington formula are not normalized in any fashion and, therefore, are not truly comparable. This means the effect of different variables within the formula remains unclear. On the other hand, Yunker and Covey, "An Optimizing Approach," p.214 (see table 4 B) avoid this pitfall by reducing the variables to percentages before utilizing them within the allocation formula they propose.
12. See, for instance, Thomas J. Pierce, "An Empirical Approach to the Allocation of the University Library Book Budget," *Collection Management* 2:47 (Spring 1978).
13. For a discussion of the pitfall of quantification, see Peter Sweetman and Paul Wiedemann, "Developing a Library Book-Fund Allocation Formula," *Journal of Academic Librarianship* 6:269 (Nov. 1980).
14. The present allocation formula uses the log base 10 to compute the log percent for selected variables. Log percent decreases the slope of the line for a plotted variable (see figure 3), minimizing the effect of large differences in variables (see figure 4). The use of linear or log percents involves a conscious, though somewhat arbitrary, choice of the mathematical approach. Alternatively, regression analysis on the raw data could be used to choose the mathematical model which would give the "best fit" for the variables being used.

15. Sweetman and Wiedemann, "Developing a Library Book-Fund," p.275.
16. See, for instance, Donna M. Goehner, "Allocation by Formula: The Rationale from an Institutional Perspective," *Collection Management* 5:165-67 (Fall/Winter 1983); S. K. Goyal, "Allocation of Library Funds to Different Departments of a University—An Operational Research Approach," *College & Research Libraries*, 34:219-22 (May 1973); Joseph A. Kohut, "Allocating the Book Budget: A Model," *College & Research Libraries* 35:192-99 (May 1974); James A. Yunker and Carol G. Covey, "An Optimizing Approach to the Problem of Interdepartmental Allocation of the Library Materials Budget," *Library Acquisitions: Practice and Theory* 4:199-223 (1980); and Packer, "Acquisitions Allocations," p.284.
17. Robin B. Devin and Martha Kellogg, "The Serial/Monograph Ratio in Research Libraries: Budgeting in Light of Citation Studies," *College & Research Libraries* 51:46-54 (Jan. 1990). For a discussion of the fallacy *pro hoc, propter hoc* and its effects, see David Hackett Fischer, *Historians' Fallacies, Toward a Logic of Historical Thought* (New York: Harper, 1970) p.368.
18. Sweetman and Wiedemann, "Developing a Library Book-Fund," p.268-76, remains, in the opinion of the author, the best single discussion of the problems of selecting variables, data collection, and their use in formulas. It is worthwhile to consult also Eikhhamenor, "A Formula for Allocating," p.148-61; and William McPheron, "Quantifying the Allocation of Monograph Funds: An Instance in Practice," *College & Research Libraries* 44:116-27 (Mar. 1983).
19. Sweetman and Wiedemann, "Developing a Library Book-Fund," p.269-71; Yunker and Covey, "An Optimizing Approach," p.1991-223, discuss this issue at length focusing on the demand for use of materials and the supply, that is, cost of materials by discipline.
20. Budd and Adams, "Allocation Formulas," p.385-86; Devin and Kellogg, "The Serial/Monograph Ratio," p.46-54.
21. Budd and Adams, "Allocation Formulas," p.386.
22. For discussion of the use of some of these sources, see Yunker and Covey, "An Optimizing Approach," p.211-12; McGrath, "A Pragmatic Book Allocation," p.356-62; and Richard Hume Werking, "Allocating the Academic Library's Book Budget: Historical Perspectives and Current Reflections," *Journal of Academic Librarianship* 14:140-44 (July 1988).
23. See Adrian Alexander, "Serial Pricing in the International Market: Fifteen Years of FAXON Experience," *Library Administration and Management* 4:27 (Winter 1990). The FAXON Company "has been providing a detailed journal price information since 1975. Detailed price analysis based on the FAXON Serials Database appears in *Library Journal* and *The Serials Librarian*." However, these data are not comprehensive nor do they deal with the many anomalies in the database such as "bill later" titles, duplicate titles for various forms of delivery by mail, different price entries for inside and outside of country of publication, and the presence of titles that we would not normally attribute to a department or discipline—reference titles, handbooks, abstracts, indexes, and directories. The UTA computer work deals with these issues. The system also eliminates specific call number areas not related to our teaching and research mission. The missing prices in the FAXON database—due primarily to bill later and direct order titles—presented a problem. Currently, these titles are assigned the average cost for other titles in the discipline.
24. The grants index proved to be an extremely knotty problem. The debate in the library advisory committee concerning its use first involved how much weight to give it. Shortly, the discussion centered on whether to count the grants funded or their dollar value. This was ultimately resolved by adopting a small special formula to combine the total number of grants and their dollar values using the formula $N=1+\log$

(\$value/1,000). This formula would give, for some possible grant sizes in chemistry, the following values of N:

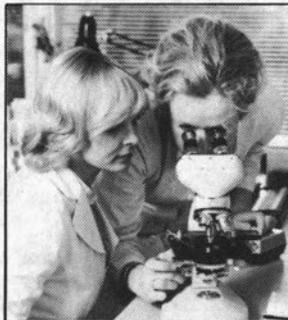
<u>Department of Chemistry</u>	
<u>Grants \$ Value</u>	<u>N</u>
\$ 2,000	1.3
10,000	2.0
100,000	3.0
1,000,000	4.0
<u>\$1,112,000</u>	<u>10.3</u>

The total N of 10.3 would then become a raw variable entry for the Department of Chemistry.

World-Class Discoveries

...begin with world-class information. It's the kind of information you'll find in *Biological Abstracts*® (*BA*) and *Biological Abstracts/RRM*® (*Reports, Reviews, Meetings*) (*BA/RRM*).

Comprehensive coverage sets *BA* and *BA/RRM* apart from other reference materials. *BA* contains full bibliographic information and abstracts from the latest biological and biomedical literature, carefully selected from approximately 7,600 journals worldwide. *BA/RRM* supplies exhaustive coverage of papers from



international meetings and symposia, as well as books, reviews and other valuable items not found in most life science reference tools. Together, *BA* and *BA/RRM* provide thorough, accurate coverage of the research that can lead to world-class discoveries!

For more information, contact BIOSIS, Marketing Department CRL392WC, 2100 Arch Street, Philadelphia, PA 19103-1399 USA. Or call toll free 1-800-523-4806 (USA except PA); (215) 587-4800 (worldwide); Telex 831739; Fax (215) 587-2016.



BIOSIS®

Information for Today's Decisions and Discoveries

BIOSIS is a registered trademark of Biological Abstracts, Inc.