The Journal of Academic Librarianship xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

The Journal of Academic Librarianship

The Effects of Reference, Instruction, Database Searches, and Ongoing Expenditures on Full-text Article Requests: An Exploratory Analysis

Jody Condit Fagan [1]

James Madison University, USA

ARTICLE INFO

Article history:
Received 20 February 2014
Accepted 2 April 2014
Available online xxxx

Keywords:
Multiple regression
Library assessment
Library metrics
ACRL Trends & Statistics Survey
Database searches
Academic libraries

ABSTRACT

ACRL Library Trends and Statistics survey data from two years were analyzed to explore relationships among variables thought to predict full-text article requests. Five different regression models predicted full-text article requests from variables including reference transactions, library instruction, Web site visits, database searches, and ongoing expenditures. Full-time student enrollment and Carnegie classification were included as covariates in all five models. All regressions supported using enrollment as a covariate; four of the regressions supported using Carnegie classification as an additional covariate. Results suggested that reference transactions, library instruction, and ongoing expenditures increase the number of database searches and full-text requests beyond these covariates. Ongoing expenditures made the largest contribution in all models, but did not always contribute the most unique variance. Furthermore, a mediation analysis showed a significant indirect effect of library instruction on full-text requests through database searches. This study suggests most ACRL data may need log transformation before using regression analyses, and offers one approach to interpreting results. Future research could investigate the meaning of ongoing expenditures after controlling for institutional size and Carnegie classification. Additional variables such as interlibrary loan usage or whether the library has a discovery tool might be interesting to include in similar models.

Academic libraries have two national sources of data about library services, expenditures, and outputs: the NCES Academic Library Survey (ALS) and the ACRL Academic Library Trends & Statistics Survey (ALTS). As library researchers have gained more familiarity with statistical techniques, the data is increasingly used on its own and with data from other sources. Strengths of these datasets are the number of institutions that participate and the number of available years of data. Thus, the sophistication of statistical analysis is not likely to be limited by sample size even after segmenting the data, dealing with missing information, or both. Having multiple years of data provides options for longitudinal studies and cross-validation of models. The two data sources can be linked with each other and with other IPEDS datasets using the IPEDS identifier, and ACRLMetrics online software (http://www.acrlmetrics.com/) permits easy selection of institutions and dataset creation and organization. For these reasons, it is important for libraries to continue to explore these existing datasets for meaning.

Research examining the influence of library data on institutional outcomes has increased in recent years. Like many areas of the academy, libraries are attempting to demonstrate their value to their institutional missions. The research agendas connecting academic libraries with student outcomes and institutional missions are promising, important, and ambitious (Hinchliffe, Oakleaf, & Malenfant, 2012; Schwieder & Hinchliffe, 2012; Stemmer & Mahan, 2012). Because so many of the relationships are indirect, however, it will take time to develop and test valid models and measures. And, the practical implications of models with such distal variables are harder to determine. Models with more proximal variables, connecting library-level inputs with library-level outputs, are still important for informing academic libraries' work and may be of interest for use as performance indicators to track institutional quality.

Electronic resources are of particular interest because of their large and increasing cost; by 2008, the average ARL library spent just over half its materials budget on electronic resources (Kyrillidou & Bland, 2009). In most cases, electronic resources entail ongoing costs with annual increases, so libraries are constantly having to re-evaluate their usefulness. Like other academic departments, libraries must demonstrate their value given the budgetary challenges facing all of higher education. In addition, preliminary research has suggested a correlation between use of electronic resources and student achievement (Goodall & Pattern, 2011).

Libraries are also concerned with the path by which users discover these expensive resources. In the age of Google, many resources can be discovered without using library services directly. However, research databases—subscription search engines usually dedicated to specific academic disciplines—are still needed by scholars and are an important

[1] School of Strategic Leadership Studies The author thanks James Madison University Libraries and Educational Technologies for access to the ACRLMetrics database.

http://dx.doi.org/10.1016/j.acalib.2014.04.002
0099-1333/

driver of electronic collection use. In addition, there is anecdotal evidence that library instruction and library reference interactions increase database and journal use through education and promotion. The library web site, with its presence on the institutional web site and library public computers, could also logically play a role in connecting users with e-resources.

The purpose of this study is to explore relationships among variables available in the ACRL Library Trends and Statistics survey, controlling for two institutional variables, full-time undergraduate and graduate student enrollment (FTE) and Carnegie classification. The influence of reference services, library instruction, web site visits, database searches, and ongoing resource expenditures on full-text requests and database searches will be the focus. While testing the hypothesized models is of some practical interest for academic libraries, learning about the relationships among these variables can also inform the development of more complex models. Structural equation modeling (SEM), for example, allows one to test theories based on data, but because so many equivalent models can fit the same dataset, SEM is best used in situations where theory has previously been developed (Kelloway, 1998; MacCallum & Austin, 2000). Because basic questions were unknown (e.g., which variables to include; whether to include covariates), it seemed prudent to learn more about relationships among the data before constructing an SEM model, especially with respect to the role of covariates.

LITERATURE REVIEW

Assessment projects related to electronic library collections and their use have historically focused on individual libraries. Examples of previously studied topics include:

- Use of electronic resources to determine curricular relevance (e.g., Kennedy, 2006; Miller, 2012).
- Correlating cost data with use data at the journal or database level to inform specific purchasing decisions (e.g., Killick, 2012).
- Using cost and use data across single-institution expenditures to provide an overview of single-institution effectiveness (e.g., Zappen, 2010).

Despite the usual focus within single institutions, there have been some studies crossing institutions using two national datasets. The NCES Academic Library Survey (ALS) has been conducted every two years since 1988 and offers descriptive statistics for approximately 3700 U.S. degree-granting institutions and their academic libraries (NCES, 2010; NCES, 2013). The primary audiences for the data are Congress and federal grant agencies, including the Institute of Museum and Library Services, National Library of Medicine, and the Library of Congress (National Center for Education Statistics (NCES), 2013). State education agencies and library administrators also use the data to plan funding, and library researchers use survey results "to determine the status of academic library operations and the librarian profession" (National Center for Education Statistics (NCES), 2011, p. 1). The NCES publishes a "First Look" report that summarizes results for each year (National Center for Education Statistics (NCES), 2010).

The ACRL Library Trends and Statistics (ALTS) survey has been conducted since 1999 and the most recent dataset (2012) included responses from 1689 libraries. There have been two reports presenting a comprehensive, historical look at the ALS data, one for the period 1974–1996 (Cahalan, Mansfield, Justh, & Mathematica Policy Research, Inc., 2001), and one covering 1994–2004 (Lu, 2007). ALS and ACRL survey data from 2000 to 2012 are available through a new subscription product called ACRLMetrics, which was described in detail by Stewart (2011, 2012). The tool offers reports for commonly desired ratios, rankings, and crosstabs, and datasets can be created and downloaded for use with statistical software.

Studies have investigated specific research topics using these two datasets, sometimes supplementing them with additional data; some of the more notable ones are reviewed here in chronological order. Regazzi (2012) analyzed U.S. academic library spending, staffing, and utilization trends from 1998 to 2008, and concluded that despite a feeling of constraint, libraries are actually receiving increases in resources over time. He also found that use of physical library assets has declined. His findings varied by type, size, and Carnegie class of institution, which suggests these institutional variables should be explored when investigating similar research questions.

Hunter and Perrett (2011) correlated ALTS data with LibQUAL+ scores across 73 universities to examine bivariate relationships between library expenditures or usage statistics and library patron satisfaction. LibQUAL+ (2013) is an instrument used by over 1000 libraries to measure satisfaction across three dimensions of library services: library as place, information control, and affect of service. The LibQUAL+ instrument measures minimum expectations, desired expectations, and perceived performance, meaning there are three scores for each of the dimensions, as well as three scores for each of the instrument's 22 items. Hunter and Perrett (2011) hoped to discover correlations that suggested which areas of library services were "the most cost-effective to fund when attempting to increase user satisfaction" (p. 407). They found some evidence the information control dimension had a strong relationship with satisfaction. Specifically, higher ACRL statistics correlated with higher patron expectations for information control, but these increases were not matched with higher perceived satisfaction scores; more expenditures and larger collections correlated with lower overall satisfaction. Their study is intriguing, but it also reveals the limitations of bivariate correlations in exploring relationships among data.

Martell (2008) examined ARL university libraries' circulation rates, reference counts, and gate counts over time and illustrated an overall decline in the use of physical collections and services and an increase in use of electronic resources. While impressive for its historical trend analysis, his study points out the importance of attempting to link inputs, which were not examined, with outputs. For example, the study did not appear to control for the change in expenditures or in volumes held over time, which could have a direct effect on circulation rates.

Yvonne Jones (2007) examined ALS data for those liberal arts colleges chosen by U. S. News and World Report as the top 50 "Best Liberal Arts Colleges." She examined the input variables Total Library Expenditure, Librarians & Other Professional Staff, Total Staff, and Books as a proportion of FTE (defined in her study as combined faculty and student FTE), and output variables Annual Circulation Transactions per FTE, Annual ILL Provided, Annual ILL Received, Reference Transactions per Week, and Gate Count per Week. She determined that $2000 was spent per FTE on library resources among the top third of best-ranked libraries; $1400 in the middle third, and $1000 for the bottom third, whereas only $200–$400 per FTE was spent by all academic libraries. She also noted that the "best-ranked" colleges had higher annual circulation transactions per FTE. Although it was unclear why the study chose to use FTE to normalize some variables but not others, her study illustrates an interesting method for using ALS data to compare libraries.

The above studies are notable for their use of the national datasets and the connections they make between library-level variables. Next, this review will examine the definitions for the specific variables from the ALTS that are of interest to this study, also noting findings from previous studies using the variables (or similar ones). The dependent variables of interest were full-text article requests and database searches. The predictor variables of interest were reference transactions, participants in group presentations, number of virtual visits to library's web site, and either ongoing electronic resource purchases (collected prior to 2012) or ongoing resource purchases (collected beginning with 2012). Full-time undergraduate and graduate student enrollment (FTE) and Carnegie classification were both used as covariates.

Please cite this article as: Fagan, J.C., The Effects of Reference, Instruction, Database Searches, and Ongoing Expenditures on Full-text Article Requests: An Exploratory Analysis, *The Journal of Academic Librarianship* (2014), http://dx.doi.org/10.1016/j.acalib.2014.04.002