8-1-2018

Understanding Public Library Services and Use: A Structural Equation Modeling Framework

Ian Burke
University of Denver

Understanding Public Library Services and Use: A Structural Equation Modeling Framework

Abstract

This dissertation describes the process of fitting publicly available data from the 2015 Public Libraries Survey and 2015 American Community Survey to a structural equation model based on a theory of action for public libraries relating operating revenue to latent factors measuring service availability and library use within the context of matched community demographic data. The process of using theory-based data analysis to investigate publicly-available data is examined. Key findings include acceptable estimates of goodness-of-fit for hypothesized factors representing public library service availability and public library use, an inconsistent mediation effect of service availability on the relationship between public library operating revenue and public library use, and identification of disparities between local community descriptors associated with library revenue and local community descriptors associated with library use. Suggestions for further investigation of these disparities are presented.

Document Type

Dissertation

Degree Name

Ph.D.

Department

Quantitative Research Methods

First Advisor

Duan Zhang, Ph.D.

Second Advisor

Denis Dumas

Third Advisor

Mary Stansbury

Keywords

American community survey, Demographics, Public data, Public libraries, Public libraries survey, Structural equation models

Subject Categories

Education | Educational Assessment, Evaluation, and Research

Understanding Public Library Services and Use:

A Structural Equation Modeling Framework.

_____

A Dissertation

Presented to

the Faculty of the Morgridge College of Education

University of Denver

_____

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

_____

by

Ian Burke

August 2018

Advisor: Duan Zhang

Author: Ian Burke
Title: Understanding Public Library Services and Use: A Structural Equation Modeling Framework.
Advisor: Duan Zhang
Degree Date: August 2018

Abstract

This dissertation describes the process of fitting publicly available data from the 2015 Public Libraries Survey and 2015 American Community Survey to a structural equation model based on a theory of action for public libraries relating operating revenue to latent factors measuring service availability and library use within the context of matched community demographic data. The process of using theory-based data analysis to investigate publicly-available data is examined. Key findings include acceptable estimates of goodness-of-fit for hypothesized factors representing public library service availability and public library use, an inconsistent mediation effect of service availability on the relationship between public library operating revenue and public library use, and identification of disparities between local community descriptors associated with library revenue and local community descriptors associated with library use. Suggestions for further investigation of these disparities are presented.

Acknowledgements

I would like to use this space to thank everyone who offered me academic and moral support throughout this process. Family and friends, I am lucky to have you all. I am also very thankful for my committee's availability and helpful feedback. I would especially like to thank my advisor, Duan Zhang, for all her guidance and patience.

Table of Contents

List of Tables

List of Figures

Chapter One: Introduction

Measuring the Modern Public Library

Public libraries exist to be used by their communities. However, the ways that libraries are used by their communities have changed over time in response to technology and community needs. In S.R. Ranganathan's text, The Five Laws of Library Science (1931), this idea is expressed as, "books are for use." This is Ranganathan's "First Law of Library Science." This sentiment is discussed in Ranganathan's text in terms of access to the printed word.

Ranganathan's "Fifth Law of Library Science," is that "the library is a growing organism." In recent decades libraries have expanded the range of their services while retaining the sentiment that the purpose of a public library – including its collections, facilities, programming, and other resources – is to be used. More than ever, librarians take pains to show that libraries offer more to the public than traditional lending services. Reported trends in library use suggest that library use is also diversifying: surveys of library users by the Pew Research Center find that a growing proportion of library users attend library programs in addition to using library media services, while a range of non-lending services including technology access and meeting space availability are now recognized as important library services (Pew Research Center, 2016).

The diversity of services offered by modern public libraries are not easy to describe succinctly. This raises the question of whether modern public library use is best measured using separate, uncorrelated indicators, or if a multivariate factor accounting for a combination of use modalities might be more apt. If this is the case, would this use factor be more effectively modeled as a single construct, or as a higher-order measurement with its own component latent measures of use? Exploring and developing methods for measuring public library use will allow for more nuanced discussion of the value of public libraries in the greater community.

In addition to the measurement of library use, this project also aims to address the measurement of library service availability. Has the expansion of library service offerings limited the extent to which one public library can be compared to another? If so, is the modern public library an organization with a singular mission, or a host for a diverse set of services with impacts that might be better measured independently? Although limited quantitative research on the relationship between library revenue, services, and use exists, the validity of the underlying models has not been extensively tested.

In the present study, I aim to use structural equation modeling techniques to construct and validate a model of public library activities using large-scale publicly-available data. Creating a structural model fill provide a framework for discussion of public library operations. In addition to testing the associational questions implied by the models tested, this project will advance discussion related to the quantitative evaluation of library services, while affording an opportunity to investigate the advantages and limitations of using large-scale administrative data to measure and evaluate public library services.

The models to be analyzed in this dissertation describe the relationship between public library operating revenue, service availability (represented as a latent combination of multiple measures), and use (also represented as a latent combination of multiple measures). The models also assess the impact of several local demographic descriptors on public library use. The models will incorporate publicly-available data from the Public Libraries Survey (PLS) and the American Community Survey (ACS). The description of these models – their fit and the estimates of relationships between library service components – will be of interest to library studies researchers, social demographers, and library administrators seeking to understand the data in these large data sets. Additionally, the process of developing this framework will provide useful information about the validity of existing theories of library service provision, equity of service provision, and library use patterns. By framing large data sets within a theory-based model, complex relationships between data points can be made more readily interpretable, while suggesting new directions for theory development and data collection procedures.

Incorporating Public Library Measurements into a Predictive Model

Public library use comes at a price: library facilities must be maintained, collections purchased, and staff compensated. The relationship between library revenue, the purchases made with that revenue, and the use of those goods and services is therefore of interest to library administrators, funding organizations, taxpayers, and the communities served by public libraries. However, existing research on this relationship using multivariate outcomes has focused on measures of materials circulation, while other studies have only considered univariate measures of public library use. This limits the extent to which existing findings can be applied to modern multi-role libraries. Incorporating the hypothetical measures described above into a model representing the ongoing work of modern public libraries will provide an example of how these measures can be applied while offering a platform for future research.

The majority of published articles using quantitative information in library studies journals consistently focus on descriptive analyses, rather than investigation of theorized causal relationships (Malliari & Togia, 2016). Analysis of quantitative relationships is more likely to be discussed in library evaluation guides, econometric studies or cost-benefit analyses published for public consumption.

These quantitative studies imply theories of action for library organizations that can be mapped to generic action models such as those proposed by (Rossi & Chen 1980; Chen, 1990) which relate organization inputs to outputs and impacts. These relationships are mediated by processes and placed in a greater environmental context. Library-specific action models have been mentioned by several sources (e.g., Shaping Outcomes, 2006; Passonneau, 2013; Sputore & Fitzgibbons, 2017). However, evidence for their use in

public library assessments is limited to the alignment of international standards for library program evaluation (ISO 11620, first adopted 2014) with categories in national library surveys such as the Public Libraries Survey (PLS). A diagram of a conceptual action model based on the theory driven/outcome-based concepts mentioned above is presented in Figure 1.

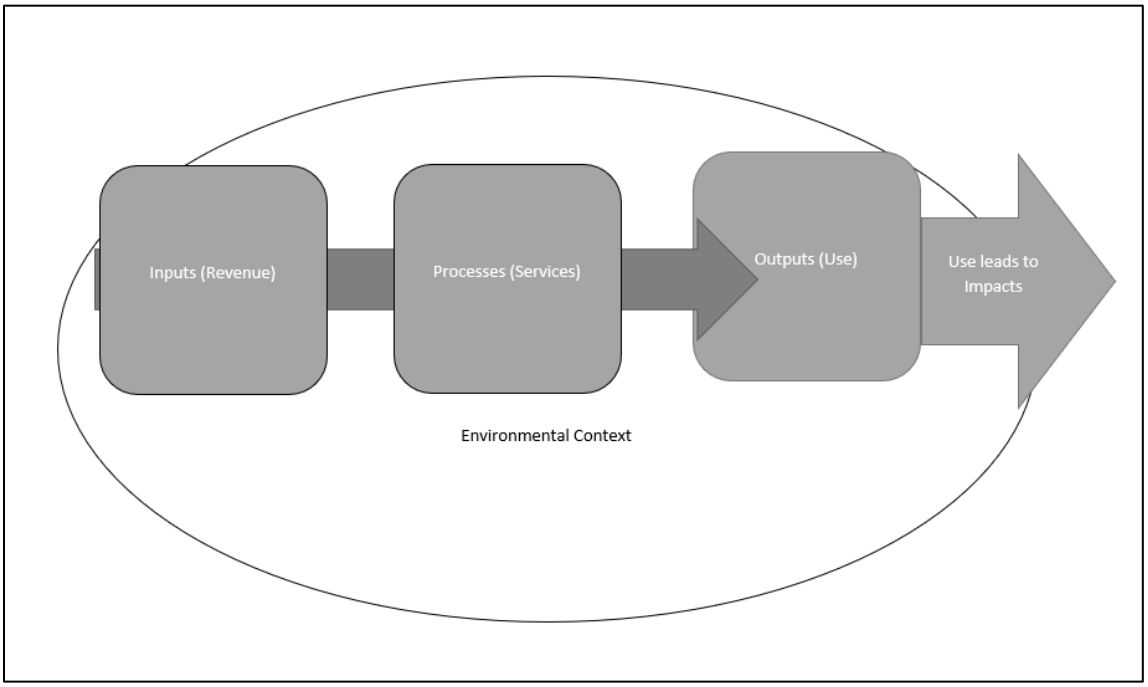

Figure 1. Conceptual diagram of an action model relating inputs to outputs, mediated by processes, within an environmental context.

Depending on the goals of the evaluation or study, the specific measures used as inputs, processes, outputs, and impacts will vary. When modeling library organizations, operating revenue might be considered an input, the use of the library can be considered an output, while the amount, quality, or type of the services provided could be all be considered mediating processes. Impacts can be defined in many ways including perceptions of value or qualitative changes in a library's community.

The models tested in the present study are informed by the theoretical foundations described above. Although theories of action for public library organizations have been explicitly described in theoretical work and implied by occasional quantitative studies, the validity of the constructs used to support these action models have yet to be examined directly. In this study, a structural equation modeling (SEM) will be used to describe the effect of library revenue (mediated by library services) and local community demographics on library revenue, services, and use. Examination of the degree to which the model fits data in the PLS data set will help to validate and/or suggest new directions for development of theories of public library action. While the model to be assessed will not include an outcome/impact analysis, it will provide a method of examining the relationship between input, process, and output variables in aggregate and in context of local demographic effects, which may be incorporated into later impact studies and library valuation efforts.

Relating Library Revenue, Services, and Use in a Structural Equation Model

Structural equation modelling (abbreviated SEM) is an applied form of the generalized linear regression model (Graham, 2008). While structural equation models equivalent to any form of regression analysis can be defined, the term "Structural Equation Modeling" is usually only applied when a model involves both directional relationships between variables and latent factors that account for measurement error within groups of related observations.

SEM is used primarily in psychometric contexts by incorporating latent factors representing psychometric concepts which in turn predict or "load" on observed indicators. There is precedent, however, for using this approach to validate theoretical models describing the operation of programs and organizations. Adedokun, Childress, and Burgess (2011) demonstrate the use of structural equations to validate program evaluation frameworks and understand the effect of the program activities (a state college internship program) versus external factors (such as students' residency status) on program outcomes (student plans to remain in-state following graduation). While the fit of the model examined provides empirical support for the program's theoretical model, Adedokun et al. also use the variance left unexplained by their model to discuss the limitations of their initial data collection procedures.

Jenatabadi and Ismail's (2014) use of structural mediation analysis to relate financial inputs to airline operations and performance also bears conceptual similarities to the methods described in this proposal by examining direct and mediated effects of system inputs on output variables. Use of SEM to study library programs and services is limited. Sin (2008) used structural equation modeling techniques to develop a macro-

level model of adolescents' decisions to use library resources. In contrast to the present study, Sin models the individual decision to use a library rather than the library's performance as a program. Other studies bearing methodological similarities to the present study include several econometric studies (Vitaliano 1997; Hemmetter, 2006; De Witte & Geys, 2013) examining the relationship between multivariate inputs to library organizations and multivariate use outputs. However, these studies identify linear combinations of observed variables rather than estimates of latent factors and focus on library lending services rather than library services in general. By examining the factor structure of current library metrics within a larger context, the present study will provide empirical evidence regarding the function of library organizations and provide a platform for further theory-building.

Placing Constructs and Predictors in Context

A theory of library action at the organization level is expressed in the international evaluation standard ISO 16439 (Methods and Procedures for Assessing the Impact of Libraries; first adopted in 2014). This standard suggests several indicators that could be considered inputs in evaluations of libraries and library services, funding, as well as staffing and existing capital resources. I have chosen to focus on operating revenue as the primary input in my structural model, while other predictors describing local demographics and indicators of existing capital resources such as collections and equipment will be included as covariates and potential moderators. Although human resources are suggested as inputs into library systems in the international standard, the allocation of staff resources might be better used as a part of the services construct described below in order to better model the availability of services offered by

libraries. There is evidence for sustained interest in the impact of funds and funding sources on public library services in research and professional literature (e.g., Holt & Elliot, 2001; Arns, Williams, & Miller, 2013).

Library processes in ISO 16439 are defined as, "activities which transform inputs into outputs" (International Organization for Standardization, 2014, p. 8). A more context-specific term for these activities might be, "library services." Examples listed in ISO 16439 include cataloguing, lending, and reference services. Although stereotypical depictions of libraries tend to focus on library services involving materials lending and bibliographic reference, public libraries in the United States offer more than these two narrow services. Programs for children and youth, internet access provision, meeting space, and media production are just some of the services offered by public libraries in the 21st century.

Orr (1968; 1973) identified six discrete library service types: document services, citation services, answer services, instruction services, facilities services and adjunct services. These services can be briefly summarized as services related to material storage and lending, services related to finding information (bibliographic reference services), non-bibliographic reference services, and the availability of library space for use, respectively. Depending on the strictness with which these definitions are applied, this typology can be adapted with little effort to modern library services that did not exist in the 20th century. Factor analysis of variables from the PLS data may align with a similar structure, or alternately suggest a simpler model containing a single latent variable representing library services.

The proposed model will use expenditures on staff, and collections, as well as library hours and program offerings as indicators of library service availability. These indicators will be predicted by a hypothesized latent factor and representing the intermediate allocation of financial resources within the library organization. This choice resembles a model suggested by Buckland (1999) in which library services are modeled as resource allocations that, in turn, predict library service use. Human resource-related measures like the number of librarians on staff, salary expenditures, and weekly hours have also been used as a proxy for the level of service provided, as in Mehra et al. (2017). The scale and composition of library service offerings are often overlooked in quantitative library research and evaluation. Including reference to these services will allow the proposed models to account for the way management decisions relate to output metrics.

In ISO 16439, library outputs are synonymous with measures of library use. these measures include loans, library visits, and program attendance. The output variable of the proposed model is a latent factor predicting library use in aggregate, similar to the latent factor model for services mentioned above. Measuring the use of libraries is understood to be an important part of librarianship, although individual public libraries take different approaches toward measurement of services. Some counts such as book and item loans are automatically collected through library circulation systems. However, other information such as library visits and reference transactions can be difficult to precisely tabulate, particularly for public library systems with limited resources. This can lead to the use of sample-based estimates rather than inventories and censuses for these metrics (Research Institute for Public Libraries Regional Proceedings, 2017). The relative

difficulty of collecting information on library use outside of materials circulation may account for a preference in prior research relating library inputs and/or services to outputs for measuring relationships for univariate outcomes or factors composed solely of circulation metrics.

The proposed model will estimate effects of selected library and local demographic characteristics on library use in order to better understand the interplay between supply and external demand on realized library use. These predictors of library use could also be said to represent the environmental context of the library system. Several significant relationships between use metrics and local community descriptors have been noted in published research. These include service population median income (Sin, 2011), service population education level (De Witte & Geys, 2013), and library funding sources (Vitaliano, 1997; Hemmeter, 2006; Sin, 2011). The percentage of households with children and percentage of population as identifying as an ethnic minority will also be included as hypothetical predictors. Including these variables in the models tested could improve their explanatory value while allowing the interaction between library funding and local characteristics to be investigated. In order to incorporate local descriptive information, the PLS data set will be matched with an external data such as the American Community Survey (ACS), an ongoing project managed by the United States Census Bureau.

In addition to placing public library operations in a demographic context, this report will also emphasize the relationship between the data collection process and the interpretation of the data. As the models examined are based on publicly-available data

sets, explaining the relationship between data collection and interpretation will further contribute to the interpretation and utilization of this data.

Research Questions

There are several large theoretical issues that I hope to address through the discussion of this project, including the application of publicly available data, the trade-offs inherent in fitting complex data to a generalized model, and the gaps between public library services and library organizations' measurement capacity. However, the four research questions stated below are more focused. These will frame the development of the measurements and models analyzed in this report.

Developing measurements. Within the broader framework of structural equation modeling, there are two methods for examining latent measurement factors that predict observations included in the measurement model: exploratory factor analysis (EFA) in which the data are classified into one or more groupings based on automated comparison; and confirmatory factor analysis (CFA) in which the data are fit to a hypothesized structure. Although multi-step methods incorporating both approaches have been used (Kline, 2011), the CFA method is used in this project to test several hypothesized measurements of library service availability and use.

Developing a measurement of public library service availability. The framing question related to confirmatory factor analysis of public library service availability could be phased the following way: **Is a factor structure measuring public library service availability supported by the PLS data set?**

There are several ways that this question might be addressed. In the present study, this is accomplished by comparting two confirmatory factor models in terms of statistical

fit and the strength with which the hypothesized use factor loads on the observed variables. As noted in the following chapter, there are several observed counts reported in the PLS data set that could be related to service availability. Some are related to the amount of money spent on the library's staff and collections, while others count non-monetary service measures such as hours open, the size of the library collection, the number of programs offered, and the number of employees within the library system.

Not all of these dimensions are addressed by multiple variables in the PLS data, which makes estimating a higher-order factor model (in which the different dimensions of service availability are grouped into more than one measurement factor) difficult. Instead, two single factor models will be compared: one including an indicator of public-internet terminal availability; the other only including indicators of staffing, programming, and services. This will allow for the discussion of the way that the availability of the library as place might be represented in data and associated limitations, along with other analysis of fit and loading scores. Other topics related to this framing question that will be discussed include the contrast between monetary- and non-monetary measurements of service availability; the extent to which staff, collections, and program counts co-vary; the dimensionality of the service availability variable; and the implications of using service availability as opposed to measures of service quality or categorical descriptions of services in a model of public library action.

Developing a measurement of public library use. A similar question to the one posed above regarding service availability can be used to frame the evaluation of library use in the present context: **Is a factor structure measuring public library use supported by the PLS data set?**

Use is measured in the PLS in several different ways, including counts for materials circulation, the number of interlibrary loans, estimated library attendance, program attendance, and PC use in the library. To provide symmetry with the Service Availability measure, the two models for Use to be compared will be single-factor models. One will include a PC use measure, while the other will not. Topics for discussion related to this question include relative model fit, factor loadings, comparisons between the structures seen in the Use and Service Availability measurements, and the extent to which different aspects of public library use are captured in the Public Libraries Survey.

Evaluating the mediating effect of public library service availability on use. If the measurements of public library service availability and use discussed above are found to have reasonable statistical fit and predictive reliability, they will be operationalized in a structural model representing public library activities in an environmental context. An input based on operating revenue will predict service availability and use, while the effects of local context on library use will also be incorporated into the model. As part of the model-development process, the mediating effect of public library services on the Revenue-to-Use relationship will also be examined. This aspect of the study will be guided by the following framing question: **Does service availability fully mediate the relationship between public library revenue and public library use?** While there is no

reason to believe that this is not the case, examining the question will help to further understand the model. In addition to discussing the extent of the mediation effect, discussion of this question will also address the gap between the relationships suggested by the model tested and the functional reality that the model approximates.

Predicting public library use. The last research question to be addressed in this report will be: **To what extent do library service availability and select community descriptors such as local median income, local educational attainment, local households with children, and the racial composition of the local community) predict library use when controlling for the effect of per-capita operating revenue on service availability?** To address this compound question, the regression weights for the predictors mentioned above will be described and interpreted. It is my hope that this will provide a functional example of the investigatory utility of statistical modeling while suggesting new routes for research on public library administration and library use.

Summary of research questions. The first two research questions are essentially questions related to measurement consistency. The information available through the PLS allows for the possibility for multivariate latent measures to be constructed that measure the general extent of library service provision or availability, and the use of those services, but the degree to which the variables indicating these concepts load on the hypothesized latent factors has not been assessed, to a great extent, in published literature.

Given that consistent factors or individual variables representing service availability and use can be identified, the next stage of the study will be to examine the quantitative relationship between library revenue, service availability, and use while

modeling differences between library service areas (LSAs) using community descriptors. Where the first two research questions are focused on determining the viability of estimating latent measurements through an inferential process, the third and fourth questions are primarily oriented toward describing the process of modern public library service provision and use that may be built upon in future research.

Practical Significance

What this study is not. Before writing further about the potential contributions of the present study, I feel that it is important to explicitly define what the study will not do. Although the models tested describe predictive relationships, they do not purport to assign causal significance to any of the predictors other than revenue and associated measures of expenditures loading on the Service Availability factor. The several demographic descriptors intended to provide additional context to the question of predicting use; significantly more research will be needed to support any assertions of the relationship between community descriptors and library use.

It is also worth re-stating for clarity that this study examines library use from the perspective of library organizations, not the individual users. A finding that there is a negative relationship between local median income and library use does not mean that low-income individuals within that LSA are less likely to use the library.

Finally, this study does not include any measure of realized library value. It should be considered a first step toward a model of library action that includes measures of library value and may also be presented in concert with research assigning value to library use, but the model presented should not be taken to imply a confusion of library use with that use's social impact.

Justification for the study. This research will directly benefit the field of library studies by providing a new perspective on the measurement of library service availability and use. In addition to providing a novel description of quantitative relationships between library services and use, the presentation of a partially latent structural equation model will provide evidence for or against the validity of conceptualizing modern public libraries as organizations with consistent service and use patterns across the continental United States. Findings from this research will offer a rich quantitative description of public library operations across the United States.

However, I believe that greater value of the present research project will be realized in further research building off the work presented in this report. The models tested here will only provide a high-level snapshot of library action and does not present a complete action model as outputs are not linked to impacts. However, the work presented in this report provides a foundational framework for conceptualizing, validating, and implanting a quantitative theory of action for library organizations. I hope that future work may make use of the relationships described in this work to design studies that examine changes in public library service and use patterns over time, as well as differences between library systems due to organization structure and setting.

It is also my goal that the dissertation on this project will prove accessible to readers with limited experience in analyzing structural equation models. As the proposed study will apply techniques developed outside of the library studies field to the analysis of public library data, I feel that it is imperative that its methods, results, and discussion of findings be accessible to the wider library studies community as well as to researchers already familiar with the methods employed. While I do not intend to write an

instructional guide to structural equation model design and interpretation, I will avoid jargon where possible and define method-specific terms in lay language where necessary. I also plan to emphasize the role that uncertainty plays in statistical analysis. As noted by Fischoff and Davis (2014), there is a degree of uncertainty in all empirical research. Even reliable data-collection methods are subject to issues with missing data and errors due to malfunctioning systems; theoretical models cannot incorporate unknown factors; natural variation in measured items leads to uncertainty regarding population averages and other descriptive statistics. By embracing rather than avoiding discussion of different sources of uncertainty in the data and model, I aim to improve the usefulness of the analysis by making complex statistical models more accessible.

Chapter 2: Literature Review

Public Library Revenue, Services, and Use in Context

The present study aims to examine the validity of a model of library action at the organizational level. Several authors have presented theories of library organizations. These include Ranganathan (1931), Orr (1973), as well as later writers drawing from management, evaluation, and psychometric theories (e.g, Holt & Elliot, 2001; Sin, 2009; Lyons & Lance, 2014). The specific action model to be tested has four major components that will be discussed in the following sub-sections: Revenue, Services, Use, and local demographic effects. Later sections of the literature review will explore how these factors have been related by other researchers and how measures of library use might be operationalized as measures of value.

Revenue. Public libraries in the United States have several avenues for funding, including local tax levies, grants from state and federal library organizations, grants from non-profit charities, and internal fundraising. The mixture of funding sources available to library organizations is dependent on several factors, including state library organizations, local property values, and library organizational structure (Vitaliano, 1997). While fines for overdue materials are still collected by many libraries, these fines do not represent a consistent revenue stream and are not reported as revenue in the PLS (IMLS, 2016).

Of the sources for library revenue mentioned above, local taxes are the largest source of revenue for most libraries in the United States. This revenue stream includes taxes levied directly by library administrative units as well as income from other local taxes designated for library use. In the 2014 fiscal year, 85 percent of per-capita library operating revenue came from local government sources, a slight increase over past years. State funding sources and private funding sources make up most of the remaining fifteen percent of per-capita operating revenue (IMLS, 2017). Nationally, private funding sources have grown to provide a greater percentage of operating revenue in recent years, but operating revenue from state-level sources is highly variable between states, with some libraries in Ohio receiving most of their operating revenue from state sources (Owens & Sieminski, 2007). Federally-sourced revenue makes up less than one percent of per-capita operating revenue and is usually distributed through grants to state libraries which then distribute the funds to individual library organizations. Most of the capital revenue – revenue designated for large, infrequent capital expenditures such as facility construction and renovation – also comes from local governmental sources. However, this proportion is smaller, with roughly 60 percent of capital revenue funds in the 2014 fiscal year coming from local sources, fourteen percent from state-level sources, three percent from federal sources, and 23 percent from other sources. In contrast to operating revenue, most libraries report receiving capital funds from a single source in a given fiscal year as opposed to a mixture of sources (IMLS 2017).

Research conducted by Sin (2011) suggests that *per-capita* operating revenue is related to population density, with urban areas receiving more funds. It is not clear from this research whether this is a linear trend or if this relationship is biased due to very high

per-capita revenues at major metropolitan library organizations such as the New York Public Library. Other relationships identified by Sin include lower levels of net funding in areas with lower median incomes, and higher per-capita revenue for library systems in the Northeast and Midwest United States relative to the national average. Research conducted by the American Library Association (2012) examining funding following the recession of 2008 found that while most libraries saw reduced revenues in the aftermath of the recession, urban libraries experienced the greatest relative fluctuations in their operating budgets during the economic recovery.

Recent opinion polls suggest that although there is strong evidence for public support of libraries in terms of sentiment (Zickuhr, Rainie, Purcell, & Duggan, 2013; Rassmussen Reports, 2014), public financial support for libraries is not unlimited. In part due to the limited availability of funds, state library organizations and other groups have conducted studies to estimate the overall economic impact and return on investment to taxpayers generated by public libraries. These studies, along with several scholarly assessments of public library economic efficiency are discussed later in this chapter.

Public library services. Library services described in the 1999 textbook *Introduction to Library Public Services* (Evans, Amodeo, & Carver, 1999) include reference services, library instruction, interlibrary loan, circulation services, reserve services, special collections, serial services, media services, and government information. The 2015 edition of the same textbook – now titled *Library Programs and Services: The Fundamentals* (Evans, Saponaro, Christie, & Sinwell, 2015) – mentions these services while adding sections on electronic resources, literacy courses and other programs, and computer access. These changes are reflective of the expansion and diversification of library services in response to a changing media landscape. In the PLS, these changes can be seen changes to library collection composition over time. In 2005, 90 percent of library collections were print materials; by 2014 print materials accounted for 66 percent of library collections. This change can be explained by a gradual reduction in total print materials and large increases in the size of electronic and audio collections (IMLS, 2017). The same IMLS report also notes a trend toward libraries offering more programs and providing more public access internet terminals. However, adoption of non-traditional services has not been uniform across public libraries in the United States. *Per-capita* program offerings are significantly higher in the Northeast region than in other regions of the country (IMLS, 2017). Real, Bertot, and Jaeger (2014) describe a "digital divide" between urban and rural libraries in terms of broadband access and digital literacy training, with rural libraries lagging in these indicators. Sin and Kim (2008) note that while most library services are provided at higher rates in high-income neighborhoods, the availability of public internet terminals was not affected.

Public library use. Individual use of library resources is often described by library researchers in terms of information-seeking behaviors. Most of these theories – such as Kulthau's information search process model (Kulthau, 1991) or Sonnenwald's information horizons theory (Sonnenwald, Wildemuth, & Harmon, 2001) – emphasize individual agency in the identification of information resources. Talja, Keso, and Pietilaèinen (1999) note a movement away from the "system-centered" paradigms of information science toward a focus on the individual information-seeking behaviors in the 1990s. Indeed, most recent research has taken a user-centered perspective when examining public library use. Vakkari and collaborators (Vakkari & Serola, 2011; Sin & Vakkari 2015; Vakkari, Aabø, Audunson, Huysmans, Kwon, Oomes, & Sin 2016) present evidence implying that individual perceived benefits from library use might drive actual use of the library, while noting that use preferences vary between users' countries of residence. Varheim (2014) presents qualitative links between library program attendance and generalized trust.

When use is examined at the library organization level, it is often framed as an outcome related to one or more predictors of use. Research examining national data sets such as the PLS suggests predictors of use at the library administrative unit level. Grimes, Manjarrez, Miller, Owen, and Swan (2013, 2014a, 2014b) produced several reports using PLS data for IMLS. This series of reports included hierarchical multiple regression models predicting several measures of change over time in library use (such as visitation, circulation, and program attendance), based on several predictors, including total revenue, staffing, and resources. In the 2012 report, circulation was found to be positively and significantly related to total revenue, but there was significant variation at the state

level, with the average effect size using group mean centering showing a negative value. While these studies examined data from the same surveys, most hierarchical regression statistics calculated are unique to the reporting year, preventing easy comparison between reporting years.

Joo and Cahill (2017) examine two measures of "children's use" – circulation of children's materials and children's program attendance – as predicted by expenditures, staff, and materials and LSA size. They find that both variables are predicted by staff, expenditures, and collection volume, suggesting that service availability might be a predictor of general use as well. Joo and Cahill used PLS data for their research.

Sin and Kim (2008) examine the relationship between multiple demographic factors and library use/non-use based on PLS and Current Population Survey (CPS) data. Results are difficult to interpret due to the number of covariates. However, a positive relationship between library use and individual income is identified. Later work by Sin (2009) presents a structural equation model describing the relationship between several indicators and latent variables representing adolescents' self-reported use public library resources. Findings from this study suggest that user characteristics, library characteristics, and neighborhood characteristics all affect adolescents' decisions to use library services.

Library operations and community demographics. Several studies link aspects of library operations to community demographics. Koontz, Jue, and Bishop (2009) note that library outlet closures between 1999 and 2003 were linked to economic factors such as declines in revenue, but that African American library service area residents and low-income residents were more likely to be affected by outlet closures. Sin (2011) notes that large inequalities exist between communities in terms of revenue and service provision. More recently, a Pew Research Group report (2016) suggests that certain library use types are linked to user demographics: library computers are more likely to be used by younger users, female users, and African American users, while higher income respondents and people with children were more likely than average to have had any interaction with library services. Recent library users were more common in urban areas, but lifetime interaction with library services was evenly distributed between urban, suburban, and rural respondents to the Pew survey.

Valuing Library Revenue, Services, and Use

Interest in assigning a quantitative value to the services provided by libraries has been embraced by some in the field, while criticized by others. Robert Orr (1973) observed that the attitudes regarding quantitative measurement of library service quality could be categorized into two groups: extreme resistance and moderate acceptance. Arns, Williams, and Miller (2013), note that the reticence observed by Orr in 1973 still exists in the 21st century.

Primary research assigning economic value to library services and use provide several examples of models relating inputs to library programs (such as revenue) to outputs (such as book loans and library visits). This body of research provides evidence

that demand for library goods and services varies between geographies and can be affected by factors both related and unrelated to library management while suggesting that several observed indicators of library services and use could be represented as composite factors.

Economic valuation of library services and use. Economic valuation methods are a response to the problem of communicating the value of non-market goods and services. These methods assess the monetary value of goods and benefits that cannot be directly priced and are used in the return-on-investment studies and cost-benefit analyses discussed later in this chapter.

The United States Government Accountability Office (GAO) has noted that although assigning monetary values to all costs and outcomes can ease the decision-making process, accurately assigning dollar values to social goods can be difficult (GAO, 2014). In addition to this issue, Ackerman (2008) identifies the uncertainty inherent in making assumptions about unknown outcomes as problematic. The lack of ethical judgement implicit in monetary valuation, particularly for cost-benefit analysis of public projects has also been criticized by multiple researchers (e.g., Scitovsky, 1941; Dworkin, 1980). Lyons and Lance (2014) have noted that attempts to relate financial inputs and outputs can blur the distinction between library outputs and outcomes, as understood by practitioners of outcome-based evaluation. Theorists outside of the library field (e.g., Scitovsky, 1941; Dworkin, 1980) have criticized economic valuation methods as they present monetary wealth as a valid way of describing collective outcomes. Particularly when communal goods, externalities and social equity issues are involved, this can lead to prioritization of practices that lead to high valuations rather than outcomes with high

social value. Conniff (2012) expands on this, noting that pricing non-market goods can lead to perverse incentives to "game the system" by emphasizing more highly-valued aspects of a program at the expense of actual program benefits.

When dealing with public goods and services such as those offered by libraries, a key concept underlying the valuation process is that of consumer surplus, or the added benefit offered by the public good to consumers who would otherwise need to purchase a substitute good or service on the market (Elliott, Holt, Hayden, and Holt, 2007; Kim 2011). Methods of establishing a value for consumer surplus contingent valuation and revealed preference methods, both of which have been used in studies of public libraries.

Contingent valuation. Contingent valuation involves the use of surveys to determine monetary values for public or other non-market goods (Mitchell and Carson, 1989; Diamond and Hausman, 1994). First described by Mitchell and Carson (1989), contingent valuation methods can be grouped into two general subcategories: "willingness-to-pay" and "willingness to accept" measures. Willingness-to-pay (WTP) measures seek to assess the amount respondents will pay for access to a non-market good or service while Willingness-to-accept (WTA) measures examine the monetary amount respondents would accept to go without the good or service. While nominally equivalent, these methods can lead to noticeably different results (e.g., Hanemann, 1991; Aabø, 2005) which may be related to respondents' income, the presence of adequate market substitutes, and human psychology (Hanemann, 1991).

Contingent valuation has been used frequently to price library services. In the UK, Morris, Hawkins, and Sumison (2001) used contingent valuation surveys to establish that library users would pay .67 British pounds to rent a library book, a lower figure than the

estimated cost of book purchase, cataloging, and maintenance of 1.41 pounds per loan. In the United States, this study has been used as a reference for establishing an estimate of user willingness-to-pay for loaned materials as a percentage of item price (e.g., Barron, Williams, Bajally, Arns, & Wilson, 2005).

Aabø (2005a; 2005b) compares the willingness-to-pay (WTP) and willingness-to-accept (WTA) methods of valuation, finding a higher WTA value than WTP value. Aabø posits that WTA might better represent a market value substitute than WTP in the case of library services, particularly those that are not readily-available on the open market. Aabø notes complications with contingent value methods such as the potential to give weight to "protest" choices (providing an unrealistically high or low value to express support or dissent) in contingent valuation surveys. In contrast to Aabø, Barron, Williams, Bajjaly, Arns, and Wilson (2005) suggest that WTP valuation surveys are easier for respondents to understand, leading to a higher response rate. Contingent valuation of public libraries has been recommended as standard practice by Elliot, et al. (2007), although implementation in published reports varies. Elliot et al. (2007) recommend using WTP surveys to establish a lower bound for valuations while noting that the sum of WTP valuations of individual public library goods and services may significantly exceed a general WTP valuation for the library as a single public good.

Revealed and stated preference valuation. Another method of assessing library use value, the revealed-preference method utilizes observation and indirect estimation rather than through self-report. A related approach, the stated-preference method presents survey respondents with hypothetical market decisions in order to record their preferences. However, stated-preference methods have not been applied in a library valuation context. Levin, Driscoll, and Fleeter's 2006 study of Ohio public libraries uses estimates of value derived from purchasing costs by applying discount rates to these costs. Kim's 2011 meta-analysis of revealed preference and contingent valuations of public library services suggests that the application of discount rates to market values – as opposed to use of primary research on consumer preferences – is typical, with discount rates ranging from 20% of an item's resale value to 100% of the new-item market price as a basis for determining the value of lending services. Kim (2011) also found that revealed-preference estimates lacked common criteria for estimating value, with some studies relying on prior work (occasionally contingent valuation methods) and others using observed prices as a basis for un-discounted price estimates.

While economic valuation studies of public library services – particularly studies using revealed preference methods – tend to rely on assumptions rather than empirical observation, research on valuation methods is more developed in studies focused on other topics, such as conservation ecology and public policy economics. In these fields, comparisons between contingent valuation, revealed preference, and stated preference value elicitation methods suggest that while different approaches produce different estimates of value, observed preference structures are not dependent on method used (Adamowicz, Louviere, & Williams,1994; Carson, Flores, & Meade, 2001; Cameron,

Poe, Ethier, & Schulze, 2002; de Blaeij, Florax, Rietveld, & Verhoef, 2003). This suggests both that further development of rigorous valuation techniques for public library services should be pursued, but also that no one monetary value is likely to represent a "true" valuation of a non-market good. In a meta-analysis of continent valuation studies related to road safety, de Blaej et al. (2003) note that not only do different approaches to monetary valuation produce different results, valuations are also affected by non-linear demand curves. This finding suggests that a better understanding of demand for library services relative to service availability could affect future public library valuation efforts.

Models Relating Revenue, Services, and Use

Library program evaluation theory and practice. Several areas of study relate public library revenue to services and/or use. Recently, systematic evaluation of library programs has presented a major venue for the theoretical discussion of this relationship. Program evaluation can be defined as a process that promotes best practices with a goal of "continuous program improvement " (CDC Program Performance and Evaluation Office, 2016). In contrast to other public services, which have embraced quantitative measures since the early 20th century, discussions of specific formal quantitative evaluation practices within the library studies field began in the 1970s. An early proponent of quantitative library evaluation, Robert Orr (1973) notes that due to the idealistic goals of public libraries, some librarians may see no reason to question the library service quality or efficiency. Orr suggests that nevertheless, quantitative evaluation can improve the management of library organizations. Indeed, support for evaluation of public libraries in the mid-20th century has been described as rooted in

demands for better, or at least more efficient management (Sputore and Fitzgibbons, 2017).

Tighter budgets and demands for institutional accountability during the 1990s and later decades coincide with more evidence of support for library evaluation in scholarly and professional literature (e.g., Holt, 1996; Bassman, Lecampagne, Korb, & Chute, 1998; Rudd, 2000). In part due to the limited availability of funds, state library organizations and other groups conducted studies to estimate the overall economic impact and return on investment to taxpayers generated by public libraries. The 1994-1999 St. Louis Public Library valuation study (Holt and Elliot, 1998; 2001) is often cited as an early example of a public library system undertaking a quantitative cost-benefit analysis.

During this time, there was also a gradual transition from a management-focused perspective on evaluation toward one focused on assessing libraries' value for taxpayers and other external stakeholders. In 1988, the United States government began collecting data through the PLS (PLS) to support assessment and comparison of library programs. In 1996, Zweizig, Johnson, Robbins, and Besant published a manual for evaluating library performance called the *TELL-IT Manual*. The attention-grabbing title of the manual is an acronym: TELL-IT stands for "Talk about the vision; Explore alternatives; Learn from what's happening; Let people know what happened; Integrate results with ongoing services; and Think about how it worked. " (Zweizig et al, 1996, p. 6). Published by the American Library Association (ALA), *TELL-IT* emphasizes qualitative methods supplementing descriptive summary statistics or constructed outcome measures that provide actionable feedback for library administrators while also providing information about library performance to external stakeholders.

The methods discussed in the *TELL-IT* manual bear some resemblance to the Context Input Process Product (CIPP) model developed by Stufflebeam, Guba, and Lincoln (1985; 1989). The CIPP model is a guide to evaluation planning that places evaluation in a cyclic context, with information gained during evaluation of a program's context, input, process, or product is used to actively improve the program and adapt an ongoing evaluation to changing program parameters (Stufflebeam, 2013). The theory-driven evaluation approach (Rossi and Chen, 1980; Chen, 1990) in contrast, emphasizes the development of a theory of a program's ongoing functions and a description of how proposed changes to the program may affect the inputs, processes, and outputs of a program. The two approaches to evaluation are not incompatible, but rather frame the evaluation discussion in different ways. CIPP could be said to be more concerned with the relationship between evaluation and the decision-making process, while theory-based evaluation is associated more strongly with the methodological design of evaluation, describing the relationship between inputs and outputs, as mediated by internal processes and external environmental factors (Alkin, 2013).

More recent guidance on evaluation provided by library organizations describes an approach called outcome-based evaluation (Shaping Outcomes, 2006; Lyons & Lance, 2014). Outcome-based evaluation bears some similarities to the methods put forward in the *TELL-IT Manual* but shows more influence from theory-driven evaluation and related approaches, including an emphasis on logic models that relate programmatic inputs, processes, and outputs to broader outcomes while accounting for organizational and environmental contexts. Notably, the guidance provided in the *Shaping Outcomes: Making a Difference in Libraries and Museums* (the online course developed by IMLS

and Indiana University describing outcome-based evaluation) presents the reporting process as a final rather than intermediate step in an evaluation cycle. This suggests an emphasis on the evaluation process itself, rather than on the changes effected by the evaluative process, as in the TELL-IT and CIPP models. This change might be related to a shift from management to funding as a motivating incentive for evaluation in library circles.

Two international standards, ISO 11620 (Library Performance Indicators; first adopted 1998) and ISO 16439 (Methods and Procedures for Assessing the Impact of Libraries; first adopted in 2014) suggest standardized vocabulary for discussing library performance indicators and evaluation. There is evidence for use of these standards in the United States and abroad in academic library assessments (e.g., Passonneau, 2013; Sputore & Fitzgibbons, 2017); evidence for their use in public library assessments is limited to the alignment of ISO 11620 with categories in national library surveys such as the PLS. As with outcome-based and theory-driven evaluation frameworks, ISO 16439 describes library systems in terms conducive to the creation of a logic model, although the international standard uses the term "impact" in place of "outcome" when discussing the external effects of library programs, such as improved childhood literacy rates or increased wages due to training received at a library.

Two criticisms of current library program evaluation practice recur in professional discussions. One is that programs undertaking evaluation tasks may overemphasize areas of success due to the biasing effects due to linking funds to evidence of performance. Streatfield and Markless (2011) suggest an emphasis on independent evaluation of national library programs to prevent bias from stakeholders focused on maintaining

program funding. Another criticism is that library evaluators do not adequately distinguish outputs from outcomes (Lyons & Lance, 2014). The Research Institute for Public Libraries (RIPL) also promotes an outcomes-focused evaluation model. In RIPL training programs, readers and participants are asked to consider the differences between the inputs, outputs, and outcomes for public-library-related research projects (RIPL, 2017).

Return-on-investment and cost-benefit analysis. In contrast to economic valuation studies, which focus on the process of assigning monetary value to non-monetary outputs, return-on-investment and cost-benefit approaches relate monetary inputs to output measures. Often, return-on-investment and cost-benefit approaches make use of economic valuation to assess the value of outputs. The studies discussed below bear some conceptual resemblance to the model implied by the proposed study's second research question and represent the bulk of library studies output that could be compared to the proposed study.

Although a distinction can be made between return-on-investment studies and cost-benefit analyses, the two terms are often used interchangeably. In economics and public policy studies, approaches to cost-benefit analyses revolve around the definition of economic efficiency and operationalization of said definition, while return-on-investment is considered a statistic (USGS, 2009). Kim (2011) describes the cost-benefit ratio as a ratio between gross benefit valuations and unit cost, while return-on-investment ratios compare net benefit to initial outlays. Nevertheless, the studies described as cost-benefit or return-on-investment analyses bear enough similarities that I have included them in the same section.

While library research journals do occasionally publish articles on library return-on-investment, or other cost-benefit analyses, most analysis of the relationship discussed in this section between library revenues and outputs has been published outside of library research journals as reports to funding agencies or the general public. The 1994-1999 St. Louis Public Library valuation study (Holt and Elliot, 1998; 2001) is often cited as an early example of a public library system undertaking a quantitative cost-benefit analysis. This study, funded by a grant from IMLS involved a contingent benefit survey using the willingness-to-accept method, identified a 4-to-1 benefit to tax dollar ratio, with survey respondents specifically placing value on library staff and collections. The authors of the St. Louis Public Library study later co-authored *Measuring Your Library's Value,* a guide to conducting cost-benefit analyses in public library settings.

Perhaps due to the cost of conducting library cost-benefit evaluations, many cost-benefit studies since the beginning of the 21st century have been backed by state libraries, regional libraries, and library consortia. Barron, Williams, Bajally, Arns, and Wilson (2005) presented a multifaceted evaluation of the secondary, tertiary, and quaternary economic effects of South Carolina Libraries spending on the local economy. Elliot, Holt, Hayden, and Holt (2007) caution against the use of economic impact assessments for most library systems, as this approach can overestimate the size of library benefits, unless the library can plausibly be considered an attraction for outside economic investment. Nevertheless, the method is attractive to report-writers who want to demonstrate library value in as many ways as possible.

The South Carolina report also included an estimate of patrons' estimate of the monetary value of services received (a contingent valuation). Similar state-wide studies

incorporating both secondary economic impacts and contingent valuation have been undertaken in Florida (Pooley, Harper, Neal, Lewis, Whitfield, & Scheibe, 2008; Harris, 2013). Earlier reports on Florida libraries combine the effects of secondary economic impacts to produce a single monetized return on investment. More recent reports have been clearer about model assumptions and potentially skewed samples influencing contingent value estimations. An economic impact analysis of the Nashville, Tennessee Public Library System (Nashville Area Chamber of Commerce Research Center, 2016) use the results of a contingent valuation survey of library users to weight market valuations sourced from cited price lists, and state whether impacts are based on flat rates or multiplied by the number of loans and visits. While secondary economic impacts are also described in the report, these impacts are presented separately from the weighted direct impact valuation estimates. A recent evaluation of the statewide impact of Ohio public libraries (Fleeter, 2016), while less detailed, also specifies the methods by which output measures were assigned monetary value (revealed preference, with varying discount rates based on the reusability of items and services), while avoiding discussion of secondary economic impacts.

A large cost-benefit study of public library systems in the United States conducted by the Colorado State Library's Library Research Service (Steffen, Lietzau, Lance, Rybin, & Molliconi, 2009) takes a somewhat different approach to describing local return on investment. In contrast to other studies conducted by state libraries, the Colorado study focused on calculating a local return on investment for several individual library districts rather than aggregating at the state level. A comparison of analyses show that the

calculated rate of return was highly variable depending on library district, particularly in cases where out-of-district users make use of a different library's services.

The studies described above make use of multiple methods to communicate the relationship between financial inputs and library use. Methods used to place value on library services in large-scale return-on-investment studies include contingent valuation methods, adjusted market values for replacement goods and services, as well as economic impact analysis. However, while varied returns on investment have been noted between states and library administrative units, these differences have not yet been investigated on a national scale, using multivariate outputs.

As noted above, the proposed study will not use monetary values to describe outputs. However, estimates of the direct and indirect effects of revenue on services and use can be interpreted as estimates of efficiency while analysis of the relationship between the service and use variables will contribute to discussions of relative value of library service profiles between LSAs.

Economic efficiency of public libraries. Return on investment studies can be described as assessments of economic efficiency (Kim, 2011). However, the studies discussed above avoid explicit discussion of economic efficiency. Several studies conducted by researchers outside of the library studies field address the issue of public library efficiency more directly through analysis of library economic efficiency. In broad terms, economic efficiency is achieved when costs are minimized and outputs or profits are maximized. The goals of the efficiency evaluation process can affect the interpretation of a cost-benefit analysis (Rabiyath 2010). Discussion of the ratio of inputs to outputs within a program or firm will be interpreted differently depending on whether one is taking a view of the possibilities within an organization or firm, comparing the efficiency to the possibilities across the sector if the evaluation is being undertaken to theorize optimal allocation of resources within a system.

Most recent research examining economic efficiency of public libraries discusses efficiency at the organization level. In these studies, an efficient production frontier is estimated using multivariate regression predicting outputs based on monetary inputs and/or resource allocation. Observed input/outputs ratios are compared to the estimated frontier for the observed input, resulting in an efficiency score. Further insights can be gained by correlating efficiency estimates with hypothesized determinants of inefficiency.

Vitaliano (1997) estimates library efficiency within New York State (excluding New York City), using staff hourly wage rates as inputs and a linear combination of circulation, hours of operation, and collection expansion as outputs. Vitaliano finds that the libraries examined operate at an average of 24% over least cost, with slight efficiency

gains attributed to library reliance on local tax revenue. Vitaliano's model somewhat confusingly uses hourly wages rather than staff expenditures as inputs, while considering a combination of library services and use metrics as an output variable. Hemmeter (2006) approaches the question of the relationship between library funding sources and library efficiency using nation-level data on libraries in the United States. Like Vitaliano (1997), Hemmeter uses stochastic frontier estimation as a method of estimating library efficiency. Hemmeter uses staffing expenditures, operating expenditures, and fixed capital such as library facilities as input and a combination of circulation, hours open, and visits per collection size to estimate outputs. While the variables used in Hemmeter's model are noticeably different from those used by Vitaliano, Hemmeter also finds a small but significant relationship between efficiency and locally-sourced revenue.

In a study of public library efficiency in the Flanders region of Belgium, De Witte and Geys (2013) test a Free Disposal Hull efficiency estimation model that incorporates operations, staff, and facilities expenditures as inputs and young borrowers, book, and media circulation as outputs. De Witte and Geys' model also incorporates a library service potential factor represented by the counts of library holdings, and controls for hypothesized demand predictors including the percentage of the library service population with a university degree, the population density, and the presence of library membership or borrowing fees. Results indicate that including demand predictors in the efficiency model lessens their apparent effect on efficiency while increasing overall observed efficiency estimates. This suggests that demand for library services is not due entirely to library service offerings and management practices. De Witte and Geys also find evidence that higher income levels, local control of revenue, higher population

density, and left-leaning regional political affiliations were associated with higher rates of library use.

Like the model proposed for this study, De Witte and Geys' (2013) model examines libraries as organizations, albeit in the context of analyzing economic efficiency. The use of a factor representing library services as distinct from library inputs also resembles the model proposed for this study. However, in the De Witte-Geys model, only materials circulation counts are considered as observed output variables.

Summary: Prior Work and the Proposed Study

Qualitative evaluations of numerous projects managed within public library organizations have been described in professional literature (e.g., Fisher, Durrance, & Hinton 2004; Varheim, 2014). However, with the arguable exception of some cost-benefit/return-on-investment studies, descriptions of full-library evaluation efforts appear to be absent from professional discussions outside of evaluation manuals and workshops, such as those described in Chapter 1. However, several studies described above describe relationships between measures of library service access, use, and predictors of library use. Sin and Kim (2008) found that non-white minorities in the 2000s were less likely to use libraries in the United States while higher incomes and educational attainment were as associated with increased library use. Joo and Cahill (2017) find that a strong relationship exists between the child-focused services at libraries and overall levels of use. Econometric studies focused on the efficiency of library organizations identify several additional predictors affecting the ratio of use to library inputs. Vitaliano (1997) and Hemmeter (2006), find that greater levels of local revenue lead to small gains in library use relative to the researchers' input estimates. De Witte and Geys (2013) identify

similar effects for service availability, higher income levels, regional political affiliation, local revenue control, and population density. These findings will inform the decisions about variables included as predictors of library use in the models described in the following chapter.

Chapter 3: Methods

Data Sources

Public Libraries Survey (PLS). The PLS has been administered nationally since 1988, with a pilot in 1985. This survey and associated data are often abbreviated as PLS in library studies literature. In regression studies, the letters PLS might also stand for the Partial Least Squares regression parameter estimation technique. This technique is not used in the present study, but the acronym is noted above in order to reduce confusion. The PLS has been administered by the Institute for Library and Museum Services (IMLS) since 2007 (IMLS, 2017). Although the content of the survey has changed over time, the survey's goal of collecting data on public libraries in the United States has remained consistent. Public use data files for PLS from 1992 through 2015 are available through the IMLS website at the time of this writing. Data are collected for a series of staggered 12-month reporting periods beginning in July of the previous year and ending in December of the reporting year. There is between a one and two-year delay between administration of the survey and publication of reports and data based on the survey.

The 2014 edition of the PLS included a sampling frame of 9,295 public libraries in the United States, the District of Columbia, Puerto Rico, and other territories (9,233 libraries in the 50 states and the District of Columbia) in its universe, with a unit response rate of 97.4 percent. The 2015 edition of the PLS included 9,251 public libraries in the United States and its territories (9,229 libraries in the 50 states and the District of Columbia) in its universe, with a unit response rate of 97.6 percent. 9,066 of the library administrative entities identified as active and providing data in the 2014 data set are also represented in the 2015 data set, while 9,052 active administrative entities in the 2015 data set were also active and providing data for the 2014 collection. The mean library service population reported in the 2014 data set was 34,833, rising to 34,926 in 2015. However, both data sets contain high numbers of small library administrative entities, with a positively skewed distribution of service areas; approximately 83% of the administrative entities reporting in both years consisted of a single library branch.

The PLS is conducted through a federal-state partnership. Information in the PLS is reported by state libraries through a web portal to IMLS. Timely reporting of data is tied to some grant awards, providing incentive for state libraries and similar entities to collect and provide data. The state libraries, in turn, collect requested information through direct survey of public library administrative entities. Library administrative entities may have multiple branches or be comprised of a single library outlet. Specific survey methods used by state libraries are direct surveys of public library administrative entities within the state and may include additional items not included on the PLS. Some states – but not all – compel participation by statute. Due to the variety of survey methods utilized by various state libraries, PLS staff advise caution when comparing states' data directly.

PLS staff also note that data from the District of Columbia, Hawaii, and United States territories other than Puerto Rico are sourced from single library administrative entities, making these areas inadvisable to include in a nested analysis of administrative entities.

The central PLS reporting tool contains several measures to ensure that data are entered accurately. Match reports compare newly entered information with information provided for previous surveys and highly-correlated variables provided in the new submission as checks on the accuracy of data (large changes in data provided are flagged and confirmation requested prior to final submissions. An additional check on accuracy requires annotations of the reasons for major changes, such as new library branches opening, or a large change in a use counts. At the national level, demographers from the U.S. Census Bureau provide a final check on data accuracy by reviewing raw data submissions prior to inclusion in the master data set. If data are missing for a library administrative entity or outlet, those data points are reported to IMLS as missing. Although the goal of the PLS is to be a comprehensive census of public libraries in the United States, some libraries do not report data, or leave fields missing.

Prior to public release, several steps are taken to reduce errors in the data and suppress cases that might threaten individual privacy. Salaries and other operating expense estimates are suppressed for libraries with low employee counts and replaced with missing values. Adjusted estimates of population for overlapping service area populations are created to reduce overestimation of service area populations by calculating ratios between state total service population estimates provided by the U.S. Census Bureau with the duplicated total count, then adjusting the reported counts by this ratio.

Imputed values for missing data (other than data deliberately suppressed) are estimated using various methods, depending on whether data for prior years is available. If prior year data are available, the prior year value adjusted for growth based on the change in figures reported for similar libraries is used as a basis for the imputed figure. If prior data are not available, present-year means or medians are adjusted for the library's service area population. If population data are not reported or unreliable, other methods may be used. Imputed figures are flagged and the specific imputation methods used are identified in the public data set. Imputed data were avoided in the present research to avoid attenuation of the data. this and other data cleaning decisions are discussed further in Section 3.3.

Data in the public access version of the PLS is organized into several levels of granularity. Three data files are produced for the publicly-available data set: a state summary data file, an administrative entity data file, and an outlet-level data file. These files are related through unique identifiers for administrative units and states. Additional levels of geographic hierarchy, such as counties/parishes associated with administrative

entities are also included in the public data sets. The proposed study will use the public version of the administrative entity data file for the 2015 PLS as the primary source for most variables.

American Community Survey (ACS). The other source of data to be used in the proposed study is the ACS (ACS) conducted by the United States Census. ACS data are not publicly available at the individual response level but are instead available as summary statistics for metropolitan areas, counties and states. While Census-derived population statistics linked to LSAs and local jurisdictions are available in the PLS public data file information on local population demographics is not available, necessitating the use of ACS data.

The ACS was initiated in 2005 as part of an initiative to generate more timely estimates for population, income, and other demographic data than would be possible with the decennial census. The ACS is an ongoing nationwide (the 50 United States, the District of Columbia, Puerto Rico, and other US territories) data collection project, which allows several estimates to be calculated for geographic areas, with a trade-off between immediacy and sample size. The sampling frames for the ACS are based on the master address file generated during the most recent decennial census. This means that the basic unit of the sampling frame is the household (or group housing). Sample selection is accomplished through a multi-stage process. First, addresses are sorted into representative sub-frames to prevent addresses from being sampled more than once in a five-year period. These sub-frames are then sampled using a stratified random sampling strategy with differential probabilities based on geography (low-population census areas are oversampled) and prior response rates (areas with historically higher response rates

are sampled at a lower rate). The initial survey invitation is distributed through the mail, with reminder post-cards and telephone follow-ups. Responses are collected through mail-in forms, an internet survey, or a telephone interview. Non-responders are contacted directly by field representatives for in-person interviews. Only after failure to conduct in-person interviews (or the identification of a sampled address as no longer actively occupied), is the target address abandoned as a non-response (Torrieri, 2014).

After the data are collected, data are processed by checking responses for implausible responses that might bias estimates, and missing data imputed using a method known as hot deck allocation, a single-imputation method. Raw data are also weighted in several stages to control for error due to sampling strategies and several demographic factors. These include sex, age, race, Hispanic origin, marital status, and housing unit type and population estimates for the weights based on the decennial Census. The weighted data are then used to calculate estimates for statistics at various level of geography, which are then organized into tables to allow for filtering by geography and demographic characteristics. Depending on the size of the population available for a given estimate, estimates may be calculated using the previous 5 years of data, 3 years of data, or the previous year's data collection. For example, estimates for a large area such as the state of Colorado can be reliably generated using one year's worth of data, while estimates for smaller or less densely-populated areas such as rural Morgan County may require data to be collected over longer period of time to accumulate enough responses to generate a statistically reliable estimate.

ACS estimates for various geographical areas are available through the United States Census website. Geography options for small areas such as census tracts or

townships vary depending on the ability of ACS to generate reliable estimates, but most data points are available for counties and larger regions. In the present study, the 2015 five-year ACS estimates were used to provide information about the library service areas' local demographics.

Measures and Variables

Indicators of library services. Variables loading on the proposed Public Library Services factor will include per-capita measures for staff-related expenditures, the number of library employees, the number of licensed databases available, the number of print serial subscriptions, number of library programs, number of public use computers, collection purchases, and non-staff, non-collection expenditures (a measure that includes facilities spending as well as program-related expenses, but not major capital projects). These variables are meant to reflect the amount of services offered by a public library, in contrast to the use of those services represented by the variables selected to load on the hypothesized Library Use factor. As with the Use factor, these variables do not capture all services offered by libraries, with one-on-one service offerings being a particular area lacking data points. Nevertheless, the variables in aggregate should be indicative of the general nature and quantity of different libraries' service offerings.

Staff-related expenditures reported in the PLS cover expenditures paying for staff (including both librarians and non-librarians) salary and benefits. Other counts reported in the PLS cover the number of staff with advanced degrees in librarianship, the number of staff with the title of librarian, and number of other staff.

The PLS count of library programs uses a broad definition of programs including any group-oriented event at or sponsored by the library. Programs offered as a series are

counted by installment rather than series. One-on-one services are not counted as programs.

Publicly available internet offerings might best be estimated with the PLS public use computer count. With recent increases in wireless connectivity and access, some method of measuring the capacity of wireless support at the library might be advisable. However, no such measure exists in the PLS now.

The PLS uses several measures to evaluate the size of library collections. These include counts of database and print serial subscriptions, counts of books/e-books, counts of audio resources, and counts of video resources. However, these metrics do not always reflect recent activity put into maintenance of the collections or acquisition of new material. For this reason, annual collection expenditures are used instead. Annual non-collection, non-staff operating expenditures (described in the PLS as "Other Operating Expenditures") are also included in the Services factor model to account for the money spent by libraries toward providing the "space as a service" as well as services such as makerspaces not covered in PLS metrics.

The number of total service hours reported in the PLS is based on the sum of open business hours in the past year for all library branches. While a measurement estimating the number of hours library facilities are open to the public would be a useful indicator of service availability, the service area aggregate score for hours open is difficult to interpret, as there is no way to estimate the balance of hours open between the various library outlets within a library district and those outlets' estimated share of total services and use. The PLS outlet file does contain information about individual branches' service hours, but none of the other indicators discussed above. An area for further research

(discussed in Chapter 5) could involve examining processes for library schedule design that improve service availability and service quality. However, this is beyond the scope of the present research.

Indicators of library use. Variables loading on the proposed Use factor may include per-capita measures for circulation, library visits, active registered users, reference transactions, total program attendance, computer use sessions, number of interlibrary loans provided or received, and number of wireless sessions. These metrics are reported as whole numbers in the PLS, but will be adjusted by the library's estimated service area population. While these variables do not capture the entirety of ways that community members might make use of library space and services, they do reflect a diverse array of library use practices.

Circulation is recorded in the PLS as a total figure representing the number of loans and renewals of library material of all types to library patrons (interlibrary lending is a separate measure). This figure includes electronic materials that are checked out through the library system, but not materials accessed through third-party databases. The PLS also includes two measures that represent subsets of the total circulation count: circulation of children's materials and circulation of electronic materials. These measures were considered for but not included in the proposed models.

Visits to the library are counted as a single variable, combining visits across library facilities. While many libraries have automated systems in place to track visits, libraries that do not have the resources to count all visits are asked to collect data for one week (Sunday-Saturday) without vacation days, high levels of staff vacation, or unusual

events that might inflate or deflate attendance from the library's typical experience and then multiplying that count by 52.

Reference transactions are included in PLS as a measure of interactions between library staff and patrons with the goal of finding a specific information resource in response to a question. PLS guidance and trainings emphasize that directional responses (for example, helping a patron find the restroom or the biography section) and technical support (e.g., assisting a patron with making a photocopy) not be included in this measure.

Program attendance is measured in the PLS by counting the number of people attending library programs, such as story time, a group class, or a seminar. This measure does not include one-on-one services that might be offered by libraries, such as tutoring. The measure is duplicated between sessions; a library with a small, but consistent audience for ongoing library programs might generate the same cumulative count of attendance as that generated by a library that has few ongoing programs but hosts several high-attendance events. As with circulation, two sub-counts of program attendance are provided: children's program attendance and young adults' program attendance.

There are two interlibrary loan-related measures in the PLS: loans to other libraries and loans from other libraries. For the proposed research, I will add the two interlibrary loan measures together prior to calculating per-capita figures in order to estimate overall interlibrary loan activity.

The computer use sessions measure in the PLS is a count of the number of times individuals have logged on to public workstations at the library. The other computer-use-related variable collected in the PLS is the wireless session count, which is based on the

number of wireless sessions as calculated by the library wireless service. A caveat to note about the wireless sessions metric is that this metric may also include sessions initiated by staff using personal wireless devices such as smartphones or tablets. For this reason – as well as diversity with regard to the way computer connectivity services may be offered by different library service organizations – these metrics are reported in the summary tables below, but not included in the models tested.

Observed predictors of library use.

Revenue-related variables. Revenue-related variables are derived from operating revenue figures in the PLS. Operating revenue includes revenue from local government sources, state government sources, federal government sources, and "other" sources (such as contributions to a library budget from a foundation or a Friends of the Library group). In the model proposed above, only operating revenues will be included. Capital revenues – such as taxes levied for a major renovation or other construction project – are less stable year-to-year than operating revenues and can bias efficiency estimates (Vitaliano, 1997). While examining revenue streams separately could allow consideration of whether increased local control has a relationship to more efficient use of resources, as suggested by Hemmeter (2006), considering total operating revenue as a single aggregate variable has benefits for model parsimony.

Local descriptors as predictors of library use. The ACS variables to be used in the proposed structural model will include several variables representing binary relationships in the library administrative entities related geographies (census-designated places and counties). The ratios that these variables will be based upon include the ratio of persons identifying as white (as opposed to other racial identities) compared with the overall population, the ratio of the 25-years old and older population holding a bachelor's degree or higher, the ratio of households with children less than 18 years old. Median income for the LSA will also be included as a predictor of library use. Summary statistics from the ACS can be found at factfinder.census.gov/faces/nav/jsf/pages/index.xhtml or queried through the Census file transfer protocol tool at www.census.gov/programs-surveys/acs/data/data-via-ftp.html.

Library descriptors. Descriptors of library organization style were considered as potential predictors of library use and moderators of the Services Use relationship. Variables considered included the count of library branches (including bookmobiles), as well as the percentage of revenue derived from local (as opposed to state or federal) sources. Although these variables show promise as predictors of different library service profiles, the classification of these profiles was judged to be beyond the scope of the present research.

Data Preparation

Matching ACS and Public Library Survey data. Cases in both the PLS and ACS data sets are described with geographic identifiers (GEOIDs) defined by the United States Geological Survey (USGS) and maintained by the American National Standards Institute (ANSI). However, the boundaries of LSAs examined by the PLS are not identified by unique GEOIDs and do not perfectly align with the boundaries of geographic entities surveyed through the ACS (described in www2.census.gov/geo/pdfs/reference/geodiagram.pdf). Demographic information from the ACS was matched using GEOIDs for counties and census-designated places associated with the LSAs and using the geographic area with a closer population match.

Filtering and transforming the data. To limit the effects of extraneous variables, several steps were taken to remove extreme outliers from the data. The range of LSA sizes covered by the PLS can lead to very large and very small library districts skewing the analysis. While this study is testing a theory of action that should be applicable to most libraries, libraries serving very small and very large populations have characteristics that set them apart from other library organizations. Small population libraries are more likely to operate with little or no full-time staff and adjust hours and service offerings accordingly. Additionally, staffing data for libraries with very small staff have been omitted from the public version of the PLS data to preserve respondents' privacy and confidentiality. In contrast, very large library organizations are likely to act not just as local community centers, but as regional or national cultural institutions. This may lead to their central branches attracting more visitors from outside their reported service area, placing a negative skew on the overall distribution of per-capita use

statistics that cannot be easily attributed to local factors. For the reasons described above, libraries with a service population below 1000 or greater than 1,000,000 were excluded from the analysis. Libraries that lacked staffing expenditure information due to a small number of employees, as well as libraries not flagged as meeting the Federal-State Cooperative System definition for public libraries were also excluded. In order to make the analysis of a covariance matrix – based off of unstandardized values – viable, service area total variables were alternately divided by the service area population or one one-thousandth of the service area population. Libraries in district of Columbia, Guam, American Samoa, Alaska, Hawaii, and Ohio were also excluded from the final analysis due to their remote and/or unique organizational structures. Ohio public libraries are funded primarily though the state library system, as opposed to the local revenue sources that support most public libraries in the United States. Libraries in the other United States territories of Puerto Rico, Northern Mariana Islands, and the U.S. Virgin Islands did not participate in the 2015 PLS.

Software Used

The primary piece of software used in the analysis of factor models and structural equations was MPlus Version 7. The method of estimation used was Robust Maximum Likelihood specifying a complex model type to correct for clustering within states. Additional correlation analysis and descriptive statistics were obtained using SPSS Version 11, while data merging and text file preparation were accomplished using Microsoft Excel 2016.

Analytic Strategies

Analytic approaches by research question.

Addressing Research Question 1: Factor structure of service availability.

It is something of a tradition in measurement model evaluation to propose and compare several approaches to the estimation of a latent measurement factor. However, the number of viable indicators for a service availability factor within the PLS data set are limited, leading to a limited number of possible variations, even if difficult-to-interpret indicators such as the average hours open per week are included. The primary measurement model proposed and tested is a single-factor model with five indicators (all natural logarithms of the base data): per-capita staff expenditures (staffexp), per-capita collection expenditures (colexp), librarians per thousand people in the service area (libr), non-librarian staff per thousand people in the service area (othemp), and library programs offered per thousand people in the service area (prog).

The proposed models include variables that reflect both the money spent on library services as well as the range of services provided by public libraries. The theory underlying this model is that the resources provided by libraries – including the

collection, programs, and support services provided by librarians and other staff – tend to covary. In other words, a library organization providing more programs than average is likely to also have more staff than average. The alternative model includes a sixth indicator, the number of internet-connected workstations available per thousand people in the service area (pc), in order to test the viability of including a measurement that addresses the dimension of using library facilities outside the context of scheduled events. This model also provides a point of comparison with the primary model. The implications of a comparison between the first model proposed in this section and this alternative are not drastic, at least compared to a model that hypothesizes multiple factors loading on a higher-order service availability factor. However, the comparison between models may suggest some directions for future investigations into the dimensionality of public library service availability. Future work (discussed in Chapter 5) may also be able to provide more insight into the question of whether different categorical mixtures of library services are present, which could in turn suggest more theoretical possibilities related to the measurement of service availability. The two proposed models measuring public library service availability are shown in Figures 2 and 3. The better-fitting model was incorporated into later structural equation models used to address later research questions.

Addressing Research Question 2: Factor structure of public library use. Similar to the Service Availability factors suggested above, the number of viable indicators for a multivariate Library Use factor are limited. The primary measurement model proposed and tested is a single-factor model with the following five indicators: annual visits to libraries in the LSA per capita (visits), reference transactions per capita (ref), circulation loans per capita (circ), total incoming and outgoing interlibrary loans per capita (ill), and program attendance per capita (attend).

In the present research, the model described above is compared to a similar model with an additional PC-use sessions *per capita* indicator (pcsess) added to the single-factor model. In addition to providing a comparison model for the primary model tested, this alternative model provides some symmetry with the alternative model proposed for Research Question 1 in that it also incorporates a variable representing access to computer services. These models attempt to incorporate several dimensions of use, including the traditional lending services provided by libraries as well as programmed use within library facilities and undirected use of within-facility services represented by the pcsess variable. These measurement models are based on a theory that public library use can be measured as a single construct at the organizational level; alternative structures that could be tested in future research are described in Chapter 5. The models described above are shown in Figures 4 and 5. The better-fitting model will be incorporated into later structural equation models used to address the later research questions.

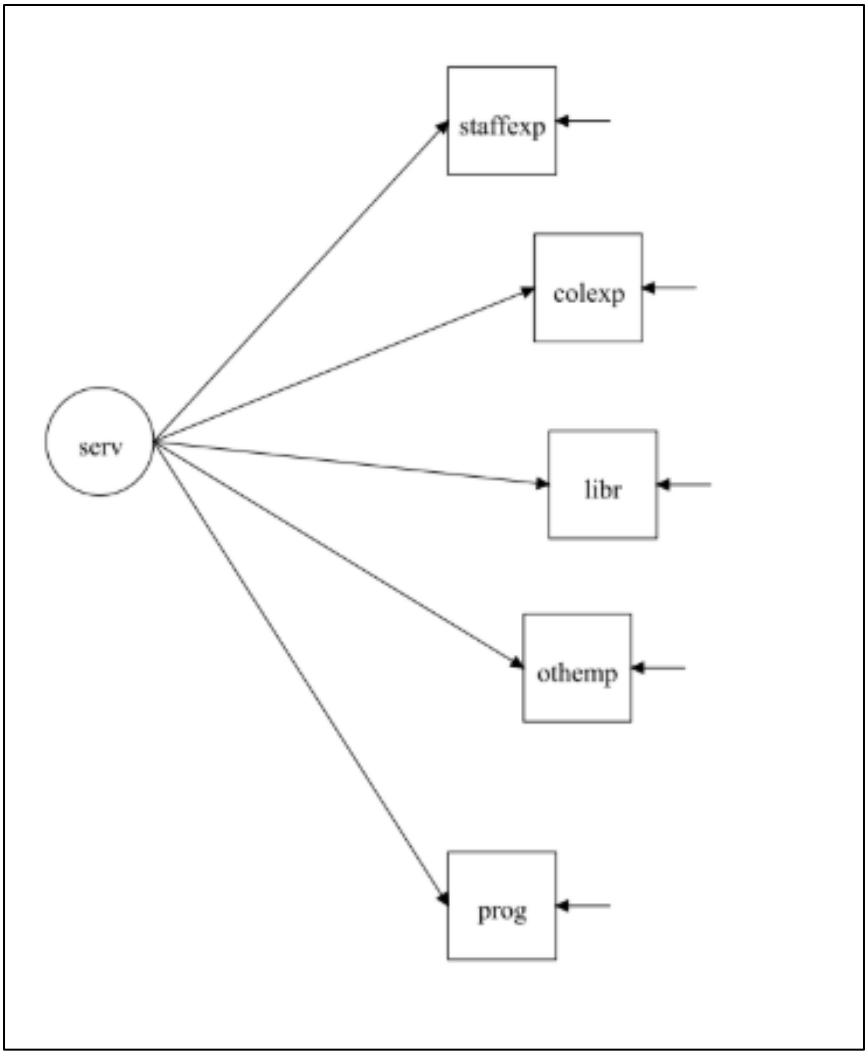

Figure 2. Five-indicator model measuring Service Availability.

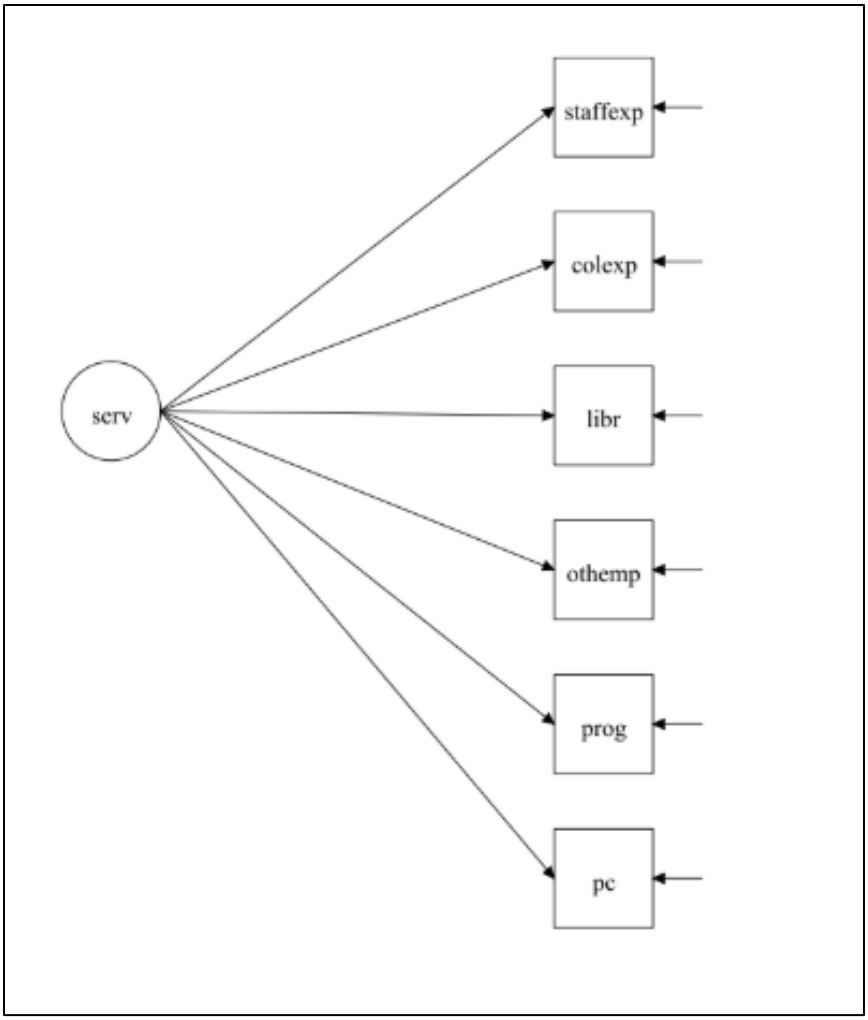

Figure 3. Six-indicator model measuring Service Availability.

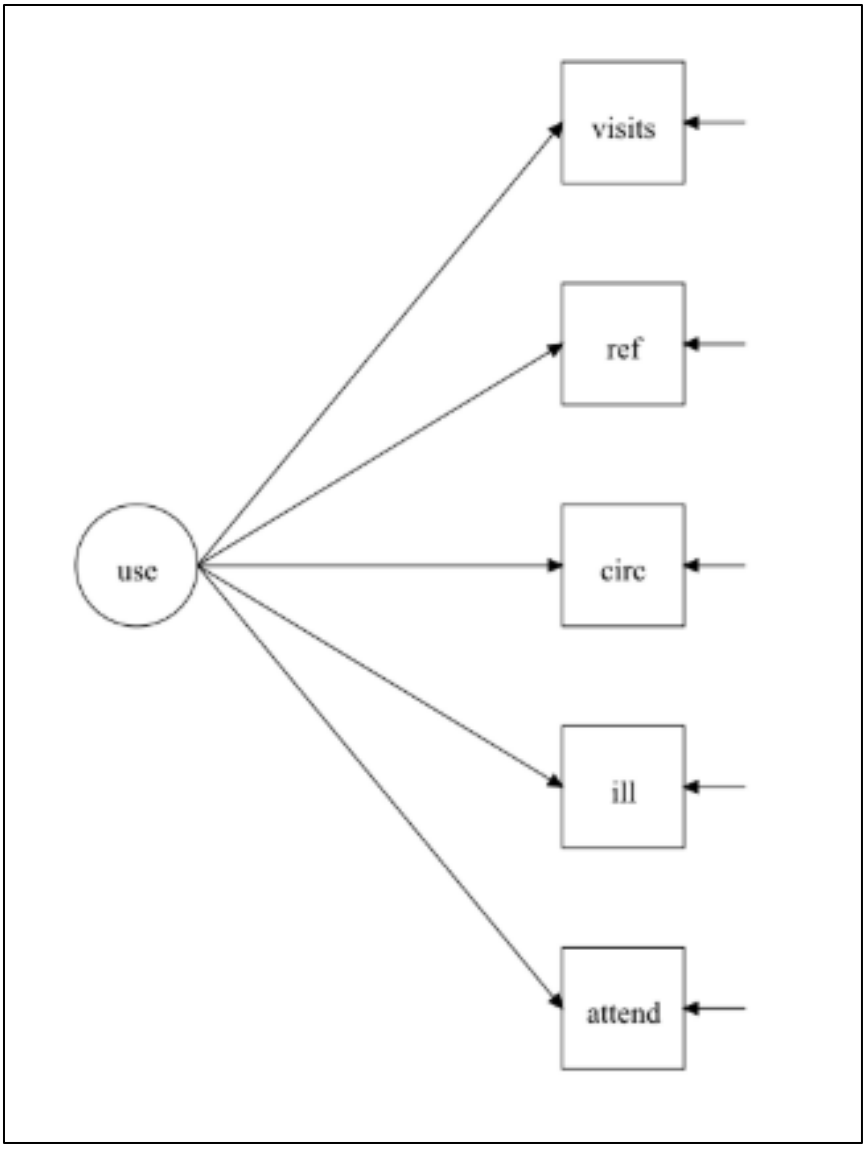

Figure 4. Five-indicator model measuring Public Library Use.

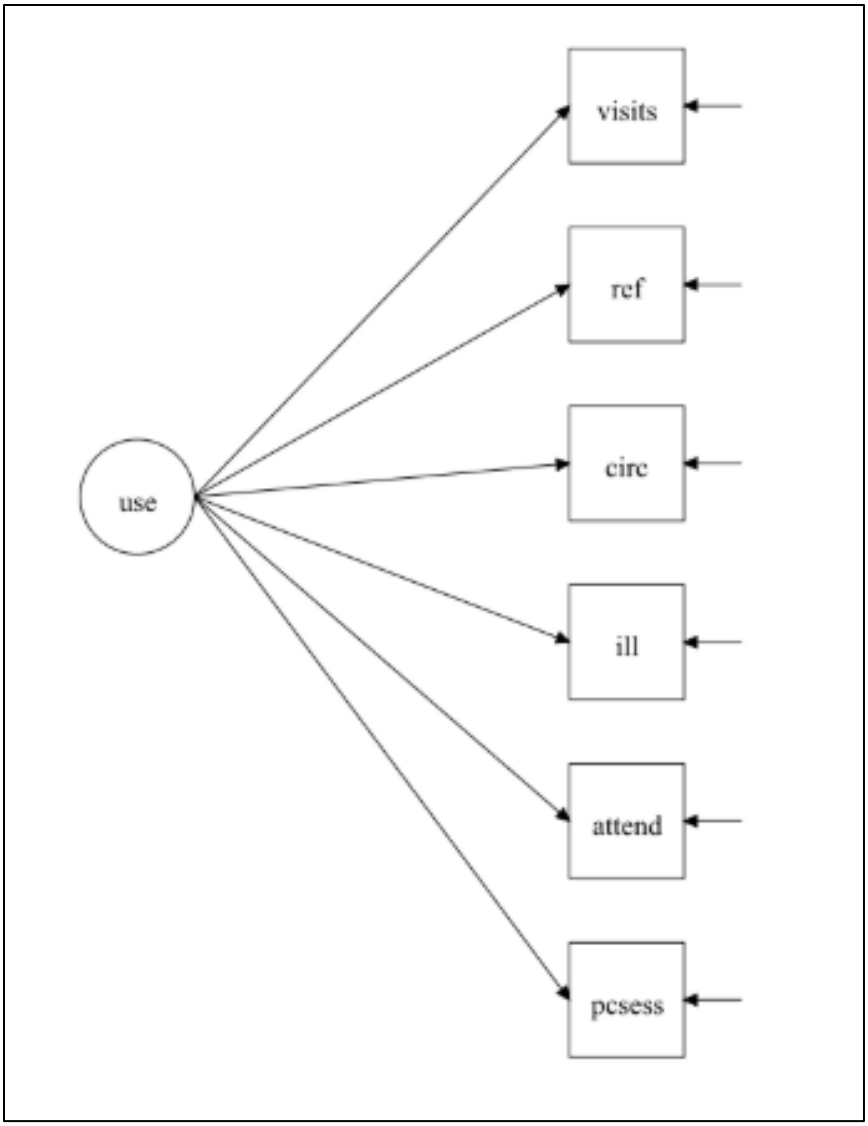

Figure 5. Six-indicator model measuring Public Library Use.

Addressing Research Question 3: Mediation of the Revenue-Use relationship.

To examine the extent to which the Service Availability factor mediates the relationship between revenue and the Library Use factor, several structural equations will be examined. To establish that Revenue, Service Availability, and Use are all related, the bivariate correlations between the Revenue variable and the Service Availability and Use factor scores will be examined and reported. To assess the nature of the mediation effect within the context of predicting library use, a structural model estimating a partial mediation of the Revenue-Use effect will be compared to a nested model that assumes complete mediation (Figure 3.3) using the Satorra-Bentler chi-square difference test (Satorra & Bentler, 2010).

Addressing Research Question 4: Prediction of public library use. While the mediating effect of the Service Availability will only be examined for the relationship between Revenue and Library Use, the direct effects of several other predictors taken from matched ACS data on the Library Use factor will also be examined. These include the percentage of persons in the area identifying as white (Race), the percentage of households with children (Child), the percentage of persons in the area over the age of 25 who graduated from college (Edu), and the median local income in thousands of dollars (Income). These variables can be seen incorporated into the models shown in Figures 6 and 7.

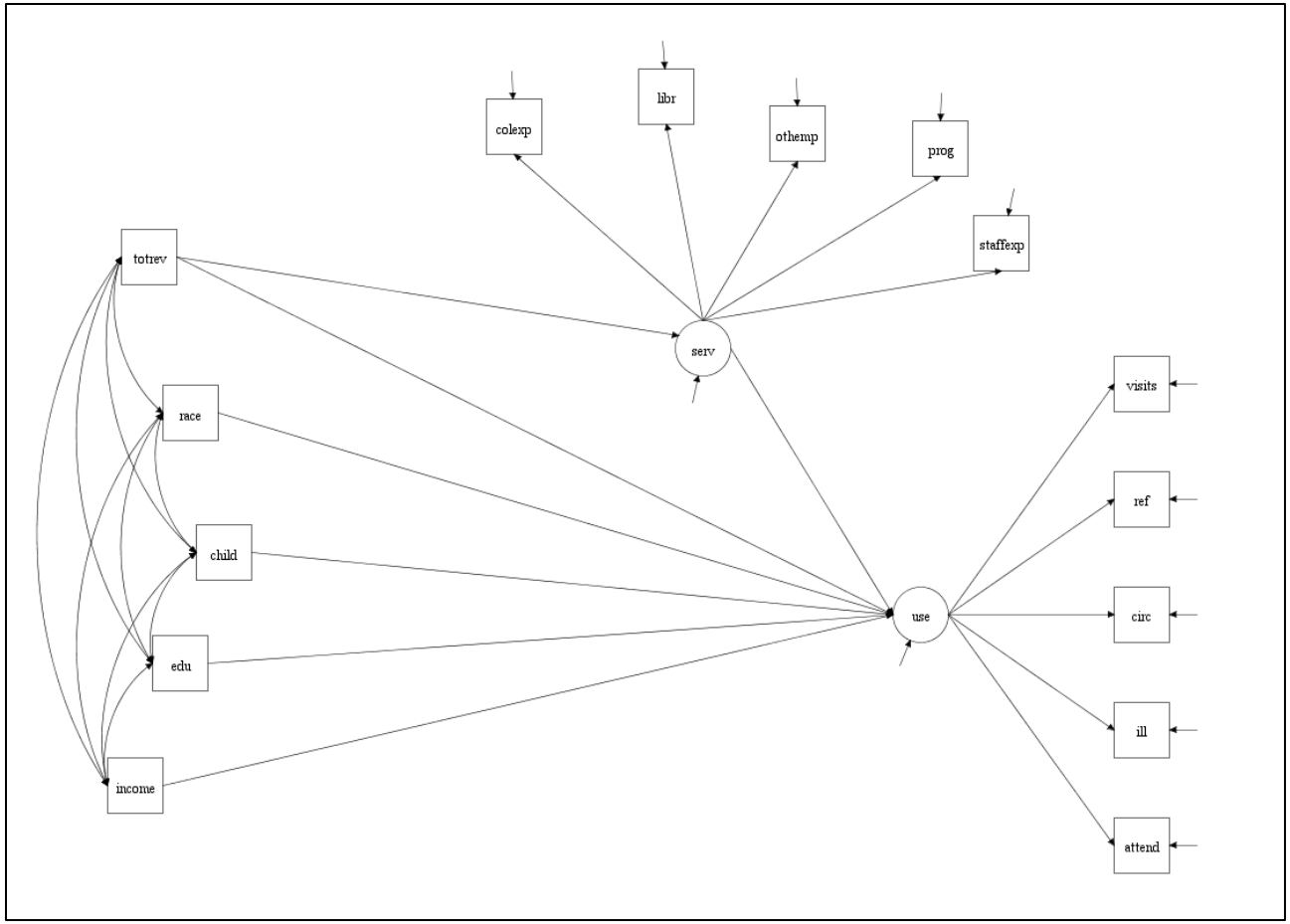

Figure 6. Structural model of library action within an environmental context assuming partial mediation of the relationship between Operating Revenue and Public Library Use by Service Availability.

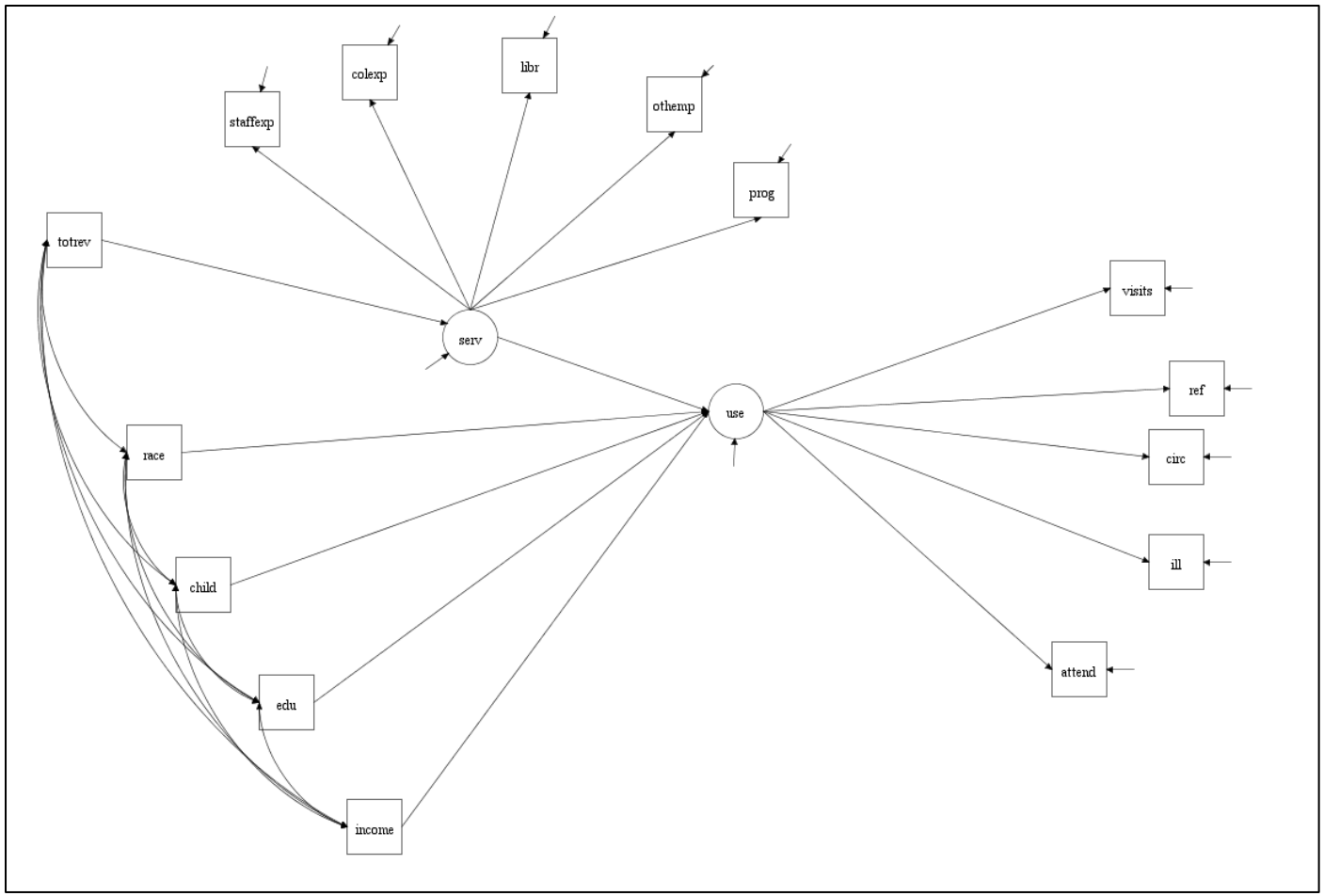

Figure 7. Structural model of library action within an environmental context assuming full mediation of the relationship between Operating Revenue and Public Library Use by Service Availability.

Statistical tests and estimated parameters.

Indices of model fit. The measurement of "fit" for confirmatory factor models and structural equation models is a common method for evaluation of the models' stability and comparing the viability of competing models. Most measures of model fit commonly used to evaluate structural models are based on the well-known chi-squared test, which compares the difference between observed values and the values expected given an assumed distribution. In the case of structural equation models, the chi-square is calculated by comparing the distribution of the matrix of residual covariances to that of an identity matrix. However, the chi-square test is sensitive to large sample sizes and is not modified by the complexity of the model. Several indices of incremental fit (in which the chi-square value of the model tested is compared to that of a "null" or "baseline" model where variables are assumed to have no correlation) and absolute fit (in which the model tested is only compared to a hypothetical complete lack of fit) have been developed to address these issues. While they make use of the chi-square statistic, they are distinct from it. David Kenny (2015) suggests a third category of fit index in addition to the incremental and absolute indices: comparative indices that cannot be assessed independently but should rather be used to compare similar models.

In addition to base chi-square statistics, the present study will report three measures of fit for the models examined: two absolute measures, one incremental measure, and one comparative measure.

One measure of absolute fit to be used is the Root Mean Square Error of Approximation or RMSEA. This measurement is calculated in the following way:

$$\sqrt{(\chi^2 - df)}/\sqrt{(df(N-1)}$$

Lower values for RMSEA are indicative of better fit than higher values. However, the question of what values for RMSEA indicate a model with unacceptably poor fit remains open. A commonly-cited rule of thumb is that a fit of less than .05 is "good," while values less than .08 are "acceptable" (McDonald & Ho, 2002; Kenny, 2015). This five-percent criterion bears some symmetry with conventions for statistical significance in hypothesis testing, potentially making the RMSEA a good choice for succinct communication of model goodness-of-fit.

The Standardized Root Mean Square Residual (SRMR) will also be reported. SRMR is another measure of absolute fit. In contrast to most other measures of fit, the SRMR does not incorporate a chi-square statistic based on the residual covariance matrix. Instead, the SRMR measures a standardized ratio of the sum of squared differences between the expected covariance and observed covariance (Hu & Bentler, 1999). As with the RMSEA, lower values are preferred, with a value of zero indicating "perfect" fit. The equation for SRMR is presented below. In this equation, s_{ij} represents the observed covariance of a pair of variables, $\sigma_{ij}^{(hat)}$ represents their covariance in the source matrix, and $s_{ii}s_{jj}$ represents the product of those variables' standard deviations. The $p(p+1)$ term represents the number of possible estimated parameters in the model, with p representing the number of observed variables included in the model.

$$\sqrt{\{2\sum_{i=1}^{p}\sum_{j=1}^{i}[\,(s_{ij} - \sigma_{ij}^{hat})/(s_{ii}s_{jj})]\text{^}2\}/p(p+1)}$$

The incremental measure to be reported is the Tucker-Lewis Index (TLI), the calculation for which is presented below. In contrast to the absolute measures of fit described above, the TLI compares the model under analysis to worst fitting model that can be estimated using the same data, referred to as the baseline or null model.

$$((\chi^2_{\text{baseline}}/df_{\text{Baseline}}) - (\chi^2_{\text{model}}/df_{\text{model}}))/((\chi^2_{\text{baseline}}/df_{\text{Baseline}}) - 1)$$

This measure is a derivate of another popular fit index known as the Comparative Fit Index (CFI), but one that penalizes model complexity more severely. In contrast to RMSEA, TLI values close to one are preferred. In general, the rule of thumb is that values greater than .9 TLI or CFI are considered acceptable fit. However, as incremental measures are based off of a comparison between the null baseline model and hypothesized models, cases in which the null model has decent absolute fit will lead to low TLI and CFI ratios (Kenny, 2015).

The Akaike Information Criterion (AIC) will also be reported as a measure of comparative or relative fit. This measure is also based on chi-square statistics, but is not a ratio, rather an adjustment to the chi-square based on the number of observations and parameters in the model:

$$\chi^2 + k(k+1) - 2df$$

This index is useful less for the purpose of evaluating goodness-of-fit but is instead used to compare non-nested models with similar goals. Thus, while it will be reported for all models analyzed, it will only be considered in cases where two or more

models are compared, as in the case of the latent factor measuring use and in the comparison of the complete and partial mediation models.

Power analysis for estimates of model fit.

The degrees of freedom for the chi-squared test are based on the difference between the number of observations – the number of unique values in the variance-covariance matrix – underlying the model and the number of parameters estimated by the model. The degrees of freedom, along with the desired alpha levels – the chance of detecting a false positive result – and the size of the sample are used to estimate the statistical power – the chance of detecting an effect if one is present – of the goodness-of-fit tests described above. Using an online power estimation calculator (Preacher and Coffman, 2006), it was determined that the sample size was more than adequate to reliably estimate model fit for all models. The minimum sample size required to identify an RMSEA of .05 with 80% statistical power for a model with five degrees of freedom is 569; the sample size of the analysis data set was 5,144 library organizations.

Average variance extracted (AVE) and composite reliability (CR). The average variance extracted (AVE) will also be reported for the factor models examined. AVE is the ratio of the squared factor loadings of a construct/factor to the sum of the squared factor loadings and residual variances in the factor's observed indicators. The composite reliability (CR) is calculated similarly, with the difference that the denominator is a squared sum, rather than a sum of squares. Unlike the measures of fit described above, these measures of construct reliability were calculated manually.

Parameter estimation. The models described above will be estimated using the robust maximum likelihood method within a single-level model adjusted for state-level clustering effects using the MPlus Type=Complex command. A comparison of multi-level and single-level adjusted models is described in Stapleton, MacNeish, and Yang (2016). Standardized and unstandardized parameters will be reported for the best-fitting structural models estimated.

In addition to estimated parameters, the Pearson correlations between the estimated Service Availability and Use factors will be reported. The primary purpose of this process will be to succinctly provide evidence for or against the hypothesis of a mediation effect of Service Availability factor on the Revenue-Use Relationship. A secondary benefit of this process is that the relationships left unanalyzed in the structural models will also be described and discussed in this report.

Satorra-Bentler scaled chi-squared difference test. To further assess the nature of the mediation relationship hypothesized, chi-square differences between the partial and full mediation models will be calculated. As the models estimated in this report are estimated using a robust maximum likelihood method, the chi-square scores will be scaled using correction factors calculated by the MPlus program during the estimation process. The calculation of these scaled differences is described by Satorra and Bentler (2010). This process increases the likelihood that a significant difference between chi-square fits will be found between nested models.

Chapter 4: Results

Summary Statistics

Geographic distribution of the data. To demonstrate some of the changes related to the filtering and data transformation processes, Table 1 compares the geographic distribution of LSAs in the complete and filtered data sets. As might be expected, the percentages for libraries and total service area populations excluded from analysis are not even across states. The filtered analysis data set is somewhat biased toward states with denser populations and/or regional public library organizations but also toward states with regional rather than local library service organizations serving rural areas.

State or State Equivalent	Count in Unfiltered Data Set	Percent of Cases in Unfiltered Data Set	Count in Filtered Data Set	Percent of Cases in Filtered Data Set	Percentage Change in Count
AK	80	0.88%	-	-	-
AL	219	2.41%	114	2.20%	-47.95%
AR	58	0.64%	48	0.90%	-17.24%
AS	1	0.01%	-	-	-
AZ	90	0.99%	40	0.80%	-55.56%
CA	184	2.03%	155	3.00%	-15.76%
CO	113	1.24%	72	1.40%	-36.28%
CT	182	2.00%	126	2.40%	-30.77%
DC	1	0.01%	-	-	-
DE	21	0.23%	21	0.40%	0.00%
FL	80	0.88%	58	1.10%	-27.50%
GA	63	0.69%	57	1.10%	-9.52%
GU	1	0.01%	-	-	-
HI	1	0.01%	-	-	-
IA	534	5.88%	162	3.10%	-69.66%
ID	102	1.12%	49	1.00%	-51.96%
IL	622	6.85%	362	7.00%	-41.80%
IN	237	2.61%	193	3.80%	-18.57%
KS	320	3.52%	116	2.30%	-63.75%
KY	119	1.31%	114	2.20%	-4.20%
LA	68	0.75%	64	1.20%	-5.88%
MA	368	4.05%	223	4.30%	-39.40%
MD	24	0.26%	23	0.40%	-4.17%
ME	228	2.51%	71	1.40%	-68.86%
MI	393	4.33%	294	5.70%	-25.19%
MN	137	1.51%	74	1.40%	-45.99%
MO	149	1.64%	105	2.00%	-29.53%
MS	52	0.57%	49	1.00%	-5.77%
MT	82	0.90%	39	0.80%	-52.44%
NC	80	0.88%	72	1.40%	-10.00%
ND	74	0.81%	16	0.30%	-78.38%
NE	253	2.79%	63	1.20%	-75.10%
NH	220	2.42%	78	1.50%	-64.55%
NJ	282	3.10%	254	4.90%	-9.93%
NM	87	0.96%	44	0.90%	-49.43%
NV	22	0.24%	17	0.30%	-22.73%
NY	756	8.32%	481	9.40%	-36.38%
OH	251	2.76%	-	-	-
OK	119	1.31%	58	1.10%	-51.26%
OR	131	1.44%	78	1.50%	-40.46%
PA	456	5.02%	317	6.20%	-30.48%
RI	48	0.53%	43	0.80%	-10.42%
SC	42	0.46%	42	0.80%	0.00%
SD	113	1.24%	30	0.60%	-73.45%
TN	186	2.05%	115	2.20%	-38.17%
TX	549	6.04%	359	7.00%	-34.61%
UT	72	0.79%	41	0.80%	-43.06%
VA	91	1.00%	81	1.60%	-10.99%
VT	159	1.75%	28	0.50%	-82.39%
WA	62	0.68%	30	0.60%	-51.61%
WI	381	4.19%	180	3.50%	-52.76%
WV	97	1.07%	65	1.30%	-32.99%
WY	23	0.25%	23	0.40%	0.00%
Total	9083	100.00%	5144	100.00%	-43.37%

Table 1. Distribution of cases by geographic entity.

Means and measures of normality. Tables 2 and 3 show changes in means and measures of normality for selected observed variables in the PLS and ACS data sets after filtering and transformation. This illustrates the changes described above. Table 4 demonstrates the effect of loglinear transformations on the scale, skewness, and kurtosis of variables in the filtered data set.

Bivariate correlations and covariances. Tables 5 and 6 show the covariance and correlation of the transformed variables described above, as well as an untransformed library service area population variable for reference. Some collinearity can be seen between revenue and indicators for the service availability and use. Moderate correlation between indicators for the Use factor and local demographic descriptors is also present. While the local descriptors related to race and households with children were not correlated with higher library revenues, local median income and college education were positively correlated with library operating revenue.

	Data Source	N	Minimum	Maximum	Mean	Std. Dev.	Skewness	Skewness SE	Kurtosis	Kurtosis SE
Library Service Area Population	PLS	9,060	19.00	4,076,438.00	34,958.75	139,566.64	14.39	0.03	295.94	0.05
Percentage of local population identifying as white	ACS	9,081	0.0%	100.0%	85.3%	0.14	-2.02	0.03	5.45	0.05
Percentage of local households with children	ACS	9,081	0.0%	80.7%	30.7%	0.08	0.26	0.03	1.72	0.05
Percentage of local population with a college degree	ACS	9,081	0.0%	90.3%	32.8%	0.14	0.93	0.03	0.88	0.05
Median Regional Household Income	ACS	9,081	0.00	$242,782	$50,183	$20,641	1.97	0.03	7.19	0.05
Total Operating Revenue per person in the Library Service Area (per capita)	PLS	9,070	0.00	15,945.11	49.53	183.90	72.81	0.03	6,178.69	0.05
Total Operating Expenditures per capita	PLS	9,070	0.00	14,162.21	46.52	163.49	72.31	0.03	6,143.92	0.05
Staff Expenditures per capita	PLS	5,934	0.00	12,738.11	33.65	169.02	71.72	0.03	5,382.99	0.06
Collection Expenditures per capita	PLS	9,069	0.00	1,361.26	5.35	15.53	74.15	0.03	6,423.14	0.05
Total Library Employees per 1000 persons in service population	PLS	9,070	0.00	157.89	0.74	1.83	71.21	0.03	6,064.70	0.05
Librarians per 1000 persons in service population	PLS	9,070	0.00	105.26	0.42	1.22	71.36	0.03	6,106.81	0.05
Non-Librarian Employees per 1000 persons in service population	PLS	9,067	0.00	52.63	0.32	0.73	44.58	0.03	2,946.04	0.05
Programs per 1000 persons in service population	PLS	9,069	0.00	3,858.70	38.41	82.87	19.27	0.03	666.92	0.05
Internet-connected public use PCs per 1000 persons in population	PLS	9,070	0.00	315.79	2.90	7.26	24.87	0.03	909.67	0.05
Visits per capita	PLS	9,070	0.00	203.47	5.93	6.00	8.06	0.03	167.55	0.05
Reference transactions per capita	PLS	9,070	0.00	96.28	0.85	1.95	22.05	0.03	843.11	0.05
Internal circulation per capita	PLS	9,070	0.00	131.87	7.80	7.14	4.32	0.03	41.64	0.05
Incoming and Outgoing interlibrary loans per capita	PLS	9,070	0.00	41.58	1.11	2.28	5.30	0.03	47.66	0.05
Program attendance per capita	PLS	9,070	0.00	87.12	0.62	1.38	35.18	0.03	1,965.64	0.05
Public PC sessions per capita	PLS	9,070	0.00	229.21	1.32	3.37	45.30	0.03	2,772.62	0.05

Table 2. Descriptive Statistics for all FSCS library administrative entities in the PLS data set.

	Data Source	N	Minimum	Maximum	Mean	Std. Dev.	Skewness	Skewness SE	Kurtosis	Kurtosis SE
Library Service Area Population	PLS	5,144	1018.00	985,310.00	43,272.54	91,224.60	5.47	0.03	37.64	0.07
Percentage of local population identifying as white	ACS	5,144	0.03	1.00	0.83	0.14	-1.42	0.03	2.50	0.07
Percentage of local households with children	ACS	5,144	0.00	0.74	0.31	0.08	0.29	0.03	1.52	0.07
Percentage of local population with a college degree	ACS	5,144	0.04	0.90	0.36	0.15	0.81	0.03	0.39	0.07
Median Regional Household Income	ACS	5,144	0.00	242,782.00	53,178.72	23,829.87	1.75	0.03	5.01	0.07
Total Operating Revenue per person in the Library Service Area (per capita)	PLS	5,144	1.51	1,944.64	48.89	56.32	12.40	0.03	312.48	0.07
Total Operating Expenditures per capita	PLS	5,144	1.24	1,689.13	45.43	46.86	11.43	0.03	317.30	0.07
Staff Expenditures per capita	PLS	5,144	0.00	644.17	30.29	29.45	5.71	0.03	68.67	0.07
Collection Expenditures per capita	PLS	5,144	0.00	190.50	4.95	5.07	11.55	0.03	357.52	0.07
Total Library Employees per 1000 persons in service population	PLS	5,144	0.05	13.78	0.66	0.53	6.47	0.03	105.08	0.07
Librarians per 1000 persons in service population	PLS	5,144	0.00	4.76	0.29	0.29	3.36	0.03	22.78	0.07
Non-Librarian Employees per 1000 persons in service population	PLS	5,144	0.00	12.16	0.36	0.38	9.53	0.03	225.43	0.07
Programs per 1000 persons in service population	PLS	5,144	0.00	1,164.63	29.93	42.29	8.27	0.03	138.98	0.07
Internet-connected public use PCs per 1000 persons in population	PLS	5,144	0.00	27.59	1.57	1.54	4.66	0.03	41.35	0.07
Visits per capita	PLS	5,144	0.00	73.46	6.05	5.09	4.08	0.03	30.43	0.07
Reference transactions per capita	PLS	5,144	0.00	96.28	0.86	2.04	28.01	0.03	1,140.84	0.07
Internal circulation per capita	PLS	5,144	0.01	131.87	8.10	6.65	4.38	0.03	51.08	0.07
Incoming and Outgoing interlibrary loans per capita	PLS	5,144	0.00	24.64	1.07	1.90	3.97	0.03	25.15	0.07
Program attendance per capita	PLS	5,144	0.00	15.94	0.56	0.74	7.23	0.03	97.62	0.07
Public PC sessions per capita	PLS	5,144	0.00	35.49	1.16	1.34	7.83	0.03	125.77	0.07

Table 3. Descriptive statistics for the filtered data set.

	N	Minimum	Maximum	Mean	Std. Dev.	Skewness	Skewness SE	Kurtosis	Kurtosis SE
Total Revenue	5,144	0.92	7.57	3.64	0.71	0.15	0.03	0.76	0.07
Total Operating Expenditures per capita	5,144	0.81	7.43	3.58	0.70	0.11	0.03	0.55	0.07
Staff Expenditures per capita	5,144	0.00	6.47	3.18	0.71	0.05	0.03	0.73	0.07
Collection Expenditures per capita	5,144	0.00	5.25	1.58	0.61	0.28	0.03	0.33	0.07
Librarians per 1000 persons in service populatio	5,144	0.00	1.75	0.24	0.19	1.75	0.03	4.53	0.07
Non-Librarian Employees per 1000 persons in se	5,144	0.00	2.58	0.29	0.20	1.79	0.03	8.54	0.07
Internet-connected public use PCs per 1000 pers	5,144	0.00	7.06	2.98	0.93	0.03	0.03	0.27	0.07
Programs per 1000 persons in service populatior	5,144	0.00	3.35	0.84	0.42	1.14	0.03	2.01	0.07
Visits per capita	5,144	0.00	4.31	1.78	0.56	0.32	0.03	0.61	0.07
Reference transactions per capita	5,144	0.00	4.58	0.50	0.41	1.82	0.03	6.45	0.07
Internal circulation per capita	5,144	0.01	4.89	2.01	0.63	-0.04	0.03	0.03	0.07
Incoming and Outgoing interlibrary loans per cap	5,144	0.00	3.24	0.50	0.60	1.24	0.03	0.93	0.07
Program attendance per capita	5,144	0.00	2.83	0.39	0.30	1.99	0.03	6.79	0.07
Public PC sessions per capita	5,144	0.00	3.60	0.68	0.38	1.56	0.03	4.35	0.07

Table 4. Descriptive statistics for transformed variables in filtered data set.

			1	2	3	4	5	6	7	8
1	Library service area population	Correlation	1	-.104**	-.102**	-.238**	.039**	.072**	0	-.095**
		Covariance	8321928461	-6724.031	-6485.593	-3057.254	278.675	975.285	-57.255	-6148.737
2	ln of Operating Revenue *per capita*	Correlation	-.104**	1	0.506	0.481	-0.001	-0.003	0.044	0.711
		Covariance	-6724.031	0.506	.971**	-0.015	-.060**	.412**	.419**	.941**
3	ln of operating expenditures *per capita* (provided for reference)	Correlation	-.102**	.971**	1	0.485	-0.001	-0.004	0.045	0.724
		Covariance	-6485.593	0.481	0.485	-0.011	-.072**	.436**	.436**	.968**
4	Percentage of local population identifying as white	Correlation	-.238**	-0.015	-0.011	1	0.02	-0.002	-0.003	-0.06
		Covariance	-3057.254	-0.001	-0.001	0.02	-.218**	-.127**	-.178**	-.031*
5	Percentage of local households with children	Correlation	.039**	-.060**	-.072**	-.218**	1	0.006	-0.001	0.049
		Covariance	278.675	-0.003	-0.004	-0.002	0.006	-.072**	.262**	-.060**
6	Percentage of local population with a college degree	Correlation	.072**	.412**	.436**	-.127**	-.072**	1	0.022	0.273
		Covariance	975.285	0.044	0.045	-0.003	-0.001	0.022	.768**	.447**
7	Median income in tens of thousands of dollars	Correlation	0	.419**	.436**	-.178**	.262**	.768**	1	5.679
		Covariance	-57.255	0.711	0.724	-0.06	0.049	0.273	5.679	.448**
8	ln of staff expenditures per thousand population	Correlation	-.095**	.941**	.968**	-.031*	-.060**	.447**	.448**	1
		Covariance	-6148.737	0.476	0.48	-0.003	-0.003	0.047	0.76	0.506
9	ln of collections expenditures per thousand population	Correlation	-.087**	.849**	.868**	.056**	-.072**	.413**	.401**	.817**
		Covariance	-4849.701	0.367	0.367	0.005	-0.003	0.037	0.581	0.353
10	ln of librarians per thousand population	Correlation	-.244**	.614**	.624**	.163**	-.115**	.170**	.171**	.603**
		Covariance	-4134.408	0.081	0.081	0.004	-0.002	0.005	0.076	0.08
11	ln of non-librarian staff per thousand population	Correlation	-.084**	.642**	.655**	-0.014	-.035*	.239**	.252**	.639**
		Covariance	-1538.215	0.091	0.091	0	-0.001	0.007	0.12	0.091
12	ln of library programs per thousand population	Correlation	-.249**	.602**	.619**	.201**	-.116**	.268**	.258**	.591**
		Covariance	-21118.956	0.398	0.401	0.026	-0.008	0.037	0.571	0.391
13	ln of PC workstations per thousand population	Correlation	-.259**	.464**	.472**	.154**	-.105**	-.088**	-.079**	.431**
		Covariance	-9872.532	0.138	0.137	0.009	-0.003	-0.005	-0.079	0.128
14	ln of visits *per capita*	Correlation	-.151**	.708**	.732**	.102**	-.120**	.368**	.293**	.712**
		Covariance	-7714.7	0.282	0.286	0.008	-0.005	0.031	0.391	0.284
15	ln of reference transactions *per capita*	Correlation	0.002	.431**	.439**	-.065**	-0.001	.224**	.216**	.435**
		Covariance	91.177	0.127	0.126	-0.004	0	0.014	0.213	0.128
16	ln of internal circulation *per capita*	Correlation	-.063**	.695**	.719**	.201**	-.122**	.481**	.377**	.701**
		Covariance	-3639.406	0.313	0.317	0.018	-0.006	0.045	0.568	0.315
17	ln of total interlibrary loans *per capita*	Correlation	-.199**	.396**	.414**	.067**	-.095**	.340**	.313**	.396**
		Covariance	-10796.719	0.168	0.172	0.006	-0.004	0.03	0.445	0.168
18	ln of program attendance *per capita*	Correlation	-.178**	.594**	.609**	.140**	-.068**	.276**	.261**	.582**
		Covariance	-4847.476	0.126	0.127	0.006	-0.002	0.012	0.186	0.124
19	ln of PC use session *per capita*	Correlation	-.087**	.478**	.485**	-0.005	-.105**	.083**	.039**	.473**
		Covariance	-3016.682	0.128	0.128	0	-0.003	0.005	0.035	0.127

*Table 5. Correlations and covariances between select observed variables in the filtered data set. N= 5,144. Pearson correlations. ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).*

			9	10	11	12	13	14	15	16	17	18	19
1	Library service area population	cor	-.087**	-.244**	-.084**	-.249**	-.259**	-.151**	0.002	-.063**	-.199**	-.178**	-.087**
		cov	-4849.7	-4134.41	-1538.22	-21119	-9872.53	-7714.7	91.177	-3639.41	-10796.7	-4847.48	-3016.68
2	ln of Operating Revenue *per capita*	cor	0.476	.614**	.642**	.602**	.464**	.708**	.431**	.695**	.396**	.594**	.478**
		cov	.849**	0.081	0.091	0.398	0.138	0.282	0.127	0.313	0.168	0.126	0.128
3	ln of operating expenditures *per capita* (provided for reference)	cor	0.48	.624**	.655**	.619**	.472**	.732**	.439**	.719**	.414**	.609**	.485**
		cov	.868**	0.081	0.091	0.401	0.137	0.286	0.126	0.317	0.172	0.127	0.128
4	Percentage of local population identifying as white	cor	-0.003	.163**	-0.014	.201**	.154**	.102**	-.065**	.201**	.067**	.140**	-0.005
		cov	.056**	0.004	0	0.026	0.009	0.008	-0.004	0.018	0.006	0.006	0
5	Percentage of local households with children	cor	-0.003	-.115**	-.035*	-.116**	-.105**	-.120**	-0.001	-.122**	-.095**	-.068**	-.105**
		cov	-.072**	-0.002	-0.001	-0.008	-0.003	-0.005	0	-0.006	-0.004	-0.002	-0.003
6	Percentage of local population with a college degree	cor	0.047	.170**	.239**	.268**	-.088**	.368**	.224**	.481**	.340**	.276**	.083**
		cov	.413**	0.005	0.007	0.037	-0.005	0.031	0.014	0.045	0.03	0.012	0.005
7	Median income in tens of thousands of dollars	cor	0.76	.171**	.252**	.258**	-.079**	.293**	.216**	.377**	.313**	.261**	.039**
		cov	.401**	0.076	0.12	0.571	-0.079	0.391	0.213	0.568	0.445	0.186	0.035
8	ln of staff expenditures per thousand population	cor	0.506	.603**	.639**	.591**	.431**	.712**	.435**	.701**	.396**	.582**	.473**
		cov	.817**	0.08	0.091	0.391	0.128	0.284	0.128	0.315	0.168	0.124	0.127
9	ln of collections expenditures per thousand population	cor	1	.590**	.581**	.557**	.433**	.683**	.392**	.740**	.384**	.561**	.431**
		cov	0.369	0.067	0.071	0.314	0.11	0.232	0.098	0.284	0.139	0.102	0.099
10	ln of librarians per thousand population	cor	.590**	1	.176**	.550**	.576**	.549**	.328**	.492**	.281**	.574**	.401**
		cov	0.067	0.035	0.007	0.095	0.045	0.057	0.025	0.058	0.031	0.032	0.028
11	ln of non-librarian staff per thousand population	cor	.581**	.176**	1	.443**	.392**	.507**	.323**	.474**	.281**	.450**	.386**
		cov	0.071	0.007	0.04	0.082	0.033	0.057	0.027	0.06	0.033	0.027	0.029
12	ln of library programs per thousand population	cor	.557**	.550**	.443**	1	.520**	.607**	.318**	.552**	.371**	.801**	.398**
		cov	0.314	0.095	0.082	0.863	0.202	0.316	0.122	0.324	0.205	0.223	0.14
13	ln of PC workstations per thousand population	cor	.433**	.576**	.392**	.520**	1	.518**	.291**	.368**	.134**	.500**	.592**
		cov	0.11	0.045	0.033	0.202	0.174	0.121	0.05	0.097	0.033	0.062	0.093
14	ln of visits *per capita*	cor	.683**	.549**	.507**	.607**	.518**	1	.474**	.765**	.395**	.636**	.619**
		cov	0.232	0.057	0.057	0.316	0.121	0.314	0.11	0.271	0.132	0.107	0.131
15	ln of reference transactions *per capita*	cor	.392**	.328**	.323**	.318**	.291**	.474**	1	.410**	.181**	.404**	.406**
		cov	0.098	0.025	0.027	0.122	0.05	0.11	0.17	0.107	0.045	0.05	0.063
16	ln of internal circulation *per capita*	cor	.740**	.492**	.474**	.552**	.368**	.765**	.410**	1	.446**	.584**	.451**
		cov	0.284	0.058	0.06	0.324	0.097	0.271	0.107	0.4	0.168	0.11	0.108
17	ln of total interlibrary loans *per capita*	cor	.384**	.281**	.281**	.371**	.134**	.395**	.181**	.446**	1	.303**	.168**
		cov	0.139	0.031	0.033	0.205	0.033	0.132	0.045	0.168	0.355	0.054	0.038
18	ln of program attendance *per capita*	cor	.561**	.574**	.450**	.801**	.500**	.636**	.404**	.584**	.303**	1	.455**
		cov	0.102	0.032	0.027	0.223	0.062	0.107	0.05	0.11	0.054	0.089	0.051
19	ln of PC use session *per capita*	cor	.431**	.401**	.386**	.398**	.592**	.619**	.406**	.451**	.168**	.455**	1
		cov	0.099	0.028	0.029	0.14	0.093	0.131	0.063	0.108	0.038	0.051	0.143

*Table 6. Correlations and covariances between select observed variables in the filtered data set, continued. N= 5,144. Pearson correlations. ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).*

Measurement and Structural Models

Models estimating Service Availability. Two models were run that estimated service availability. Service Availability Model 1 contained five indicators (all natural logarithms of the base data): per-capita staff expenditures (staffexp), per-capita collection expenditures (colexp), librarians per thousand people in the service area (libr), non-librarian staff per thousand people in the service area (othemp), and library programs offered per thousand people in the service area (prog). Service Availability Model 2 included an additional measurement of the number of public computer workstations available per thousand people in the service area (pc). Estimates of model fit and reliability for these models are presented below in Table 7, while the standardized errors and factor loadings are presented in Figures 8 and 9.

	Service Availability Model 1 (five indicators)			Service Availability Model 2 (six indicators)		
		df	p-value		df	p-value
Chi-square	126.17	5	<.001	262.29	9	<.0001
RMSEA	0.069			0.074		
SRMR	0.063			0.078		
TLI	0.947			0.841		
AIC	16504			20579.29		
AVE	0.58			0.53		
CR	0.87			0.87		
Scaling correction factor	12.3797			12.1814		

Table 7. Summaries of model fit and reliability for proposed Service Availability factors.

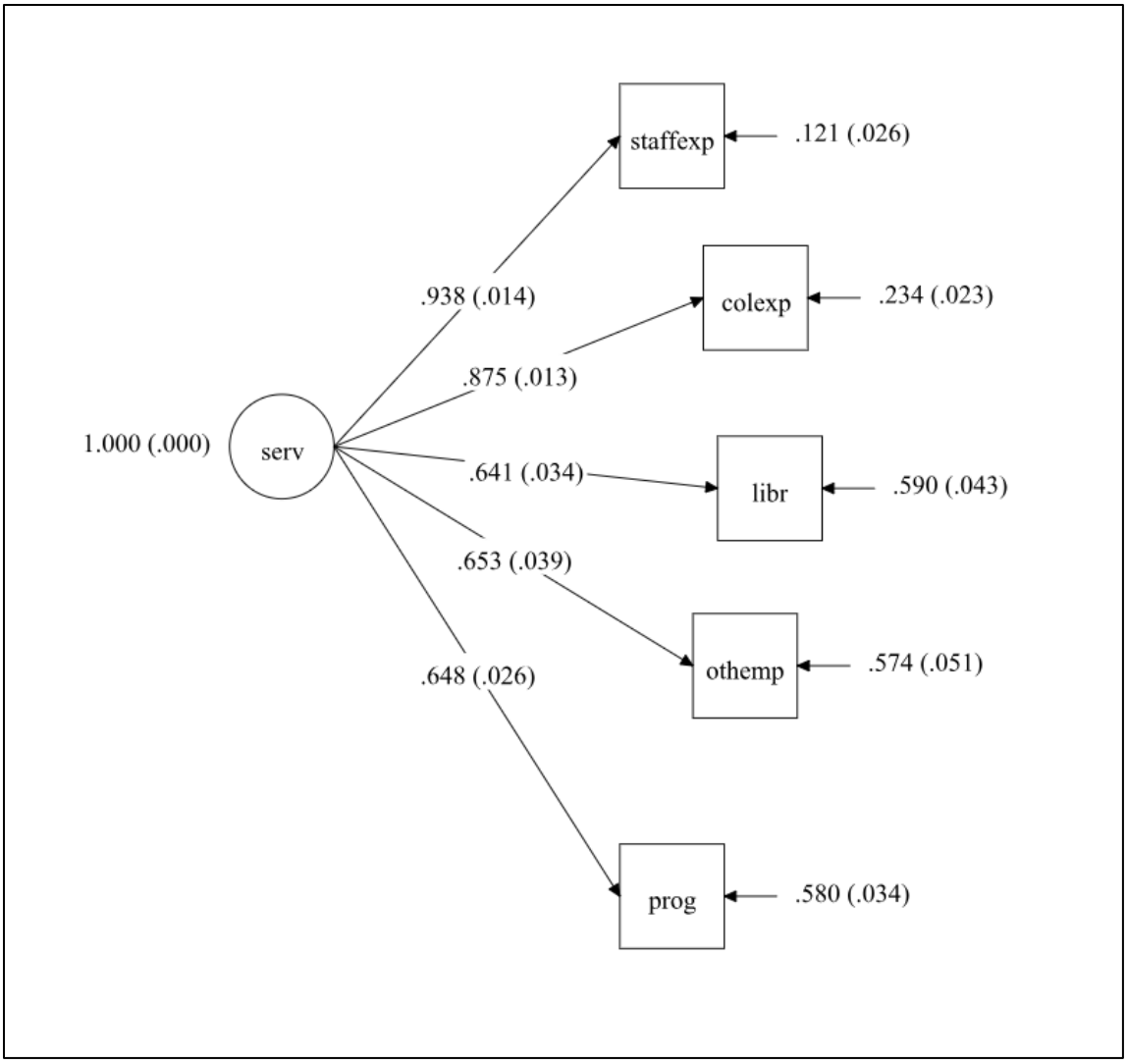

Figure 8. Standardized coefficients for a five-indicator model measuring Service Availability. Standard errors for parameter estimates are shown in parentheses.

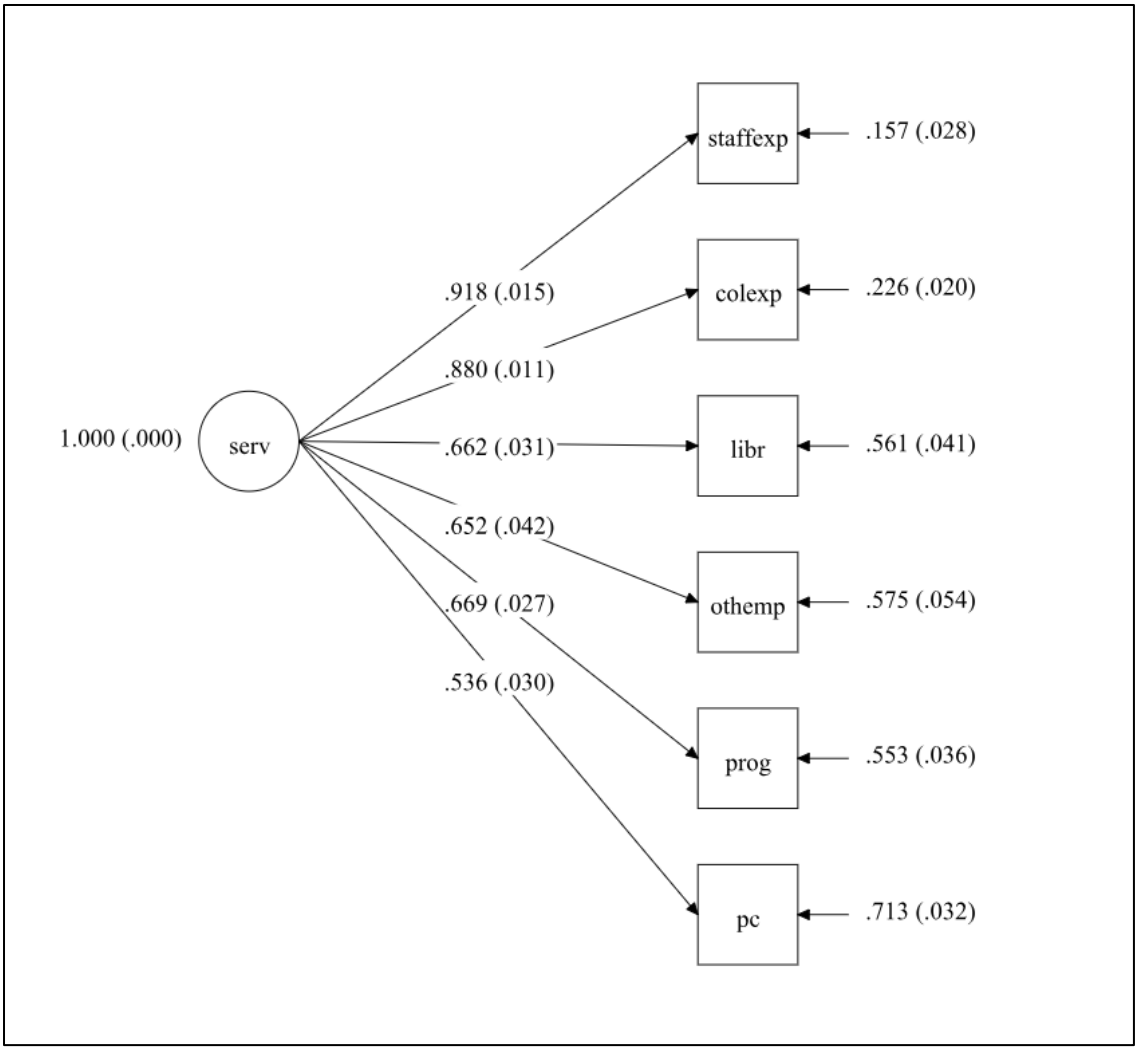

Figure 9. Standardized coefficients for a six-indicator model measuring Service Availability. Standard errors for estimates are shown in parentheses.

All five indicators for the Service Availability factor in Model 1 had loadings greater than .500, offering support for the validity of the construct. Fit indices provided limited evidence for the adequacy of model fit, with RMSEA and SRMR values below the .08 threshold and TLI value greater than .9. Model 2 performed slightly worse with regard to fit on all measures, although factor loadings remain high for the indicators.

Although the two models purporting to measure library service availability are not precisely nested, they are very similar, only differing in terms of indicators by one variable, and in degrees of freedom by four. Model fit indices and measures of reliability reflect this, with both models falling within widely-accepted ranges for "acceptable" fit on the RMSEA and SRMR indices. There is a noticeable discrepancy between TLI scores, however. Unlike RMSEA and SRMR, the TLI is an incremental fit index dependent on the average correlation between variables in the model. This suggests that the PC availability indicator is less-well correlated with the other indicators in the model than the five other indicators are with each other.

Average variance extracted (AVE) and composite reliability (CR) measures suggest that the Service Availability factor is a reliable predictor of the proposed indicators, but that less than 60 percent of the variance in the indicators is explained.

Factor loadings in both models were robust. The lowest loading was the PCs indicator in the six-factor model, with a standardized loading of .536. The two expenditure-related indicators – staff expenditures and collection expenditures – had the highest loadings in both models, with values greater than .8 for collection expenditures and values greater than .9 for staff expenditures in both models.

All measurements of model fit and reliability suggested that including the public computer variable included in the PLS data set would reduce the coherence of a single-factor measure of public library use. Re-examination of bivariate correlations between indicators reinforces this conclusion. The average Pearson correlation for the five indicators in Model 1 is .55. Adding the PC availability indicator drops the average correlation to .53. While the PC indicator has a positive absolute correlation above .5 with the program availability indicator and the number of librarians employed per thousand residents of the LSA, it has lower correlations with the expenditure-related variables as well as the transformed count non-librarian employees. The relatively low correlation between PCs and non-librarian employees is difficult to explain outside of unexamined differences in terms of library service offerings between libraries, a plausible theory regarding the low correlation between operating expenditures and PCs may be due to the infrequent nature of PC purchases. While the replacement of technology is considered an operating expense rather than a capital expense according to the official PLS documentation, the expected multi-year lifespan of personal computers could lead to these services having less of an impact on library finances than staffing, program offerings, or collection maintenance.

While Model 1 may be limited with regard to the extent that it approximates all dimensions of library service provision, its relative superiority with regard to fit and rough equivalence in terms of reliability statistics (AVE and CR) when compared to Model 2 led to its use in the structural models tested later in the project.

Models estimating Use. Two models were also run estimating a single factor representing library use. Use Model 1 includes five indicators: annual visits to libraries in the LSA per capita (visits), reference transactions per capita (ref), circulation loans per capita (circ), total incoming and outgoing interlibrary loans per capita (ill), and program attendance per-capita (attend).

As with the hypothesized Service Availability factor, Use Model 2 added an additional measurement related to internet access and computer use. Model 2 adds a sixth indicator resenting the number of public computer use sessions per thousand people in the service area (PCSESS).

Estimates of model fit and reliability for these models are presented below in Table 8, while standardized errors and factor loadings are presented in Figures 10 and 11.

	Use Model 1 (five indicators)			Use Model 2 (six indicators)		
		df	p-value		df	p-value
Chi-square	21.742	5	0.0006	97.508	9	<.0001
RMSEA	0.026			0.044		
SRMR	0.022			0.037		
TLI	0.994			0.977		
AIC	25609.17			27865.97		
AVE	0.5			0.48		
CR	0.82			0.81		
Scaling correction factor	7.4302			5.779		

Table 8. Summaries of model fit and reliability for proposed Use factors.

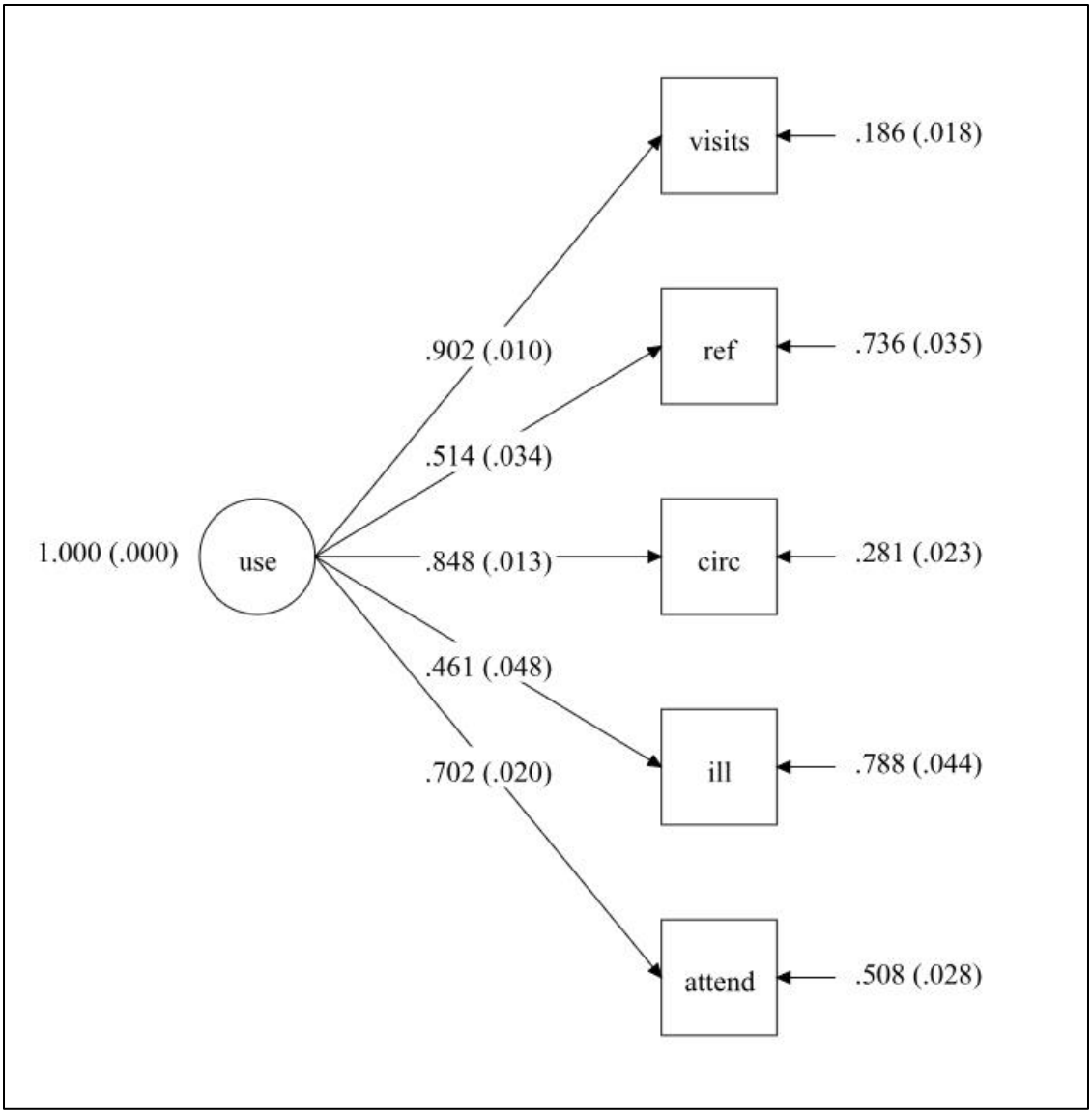

Figure 10. Standardized coefficients for a six-indicator model measuring Public Library Use. Standard errors for parameter estimates are shown in parentheses.

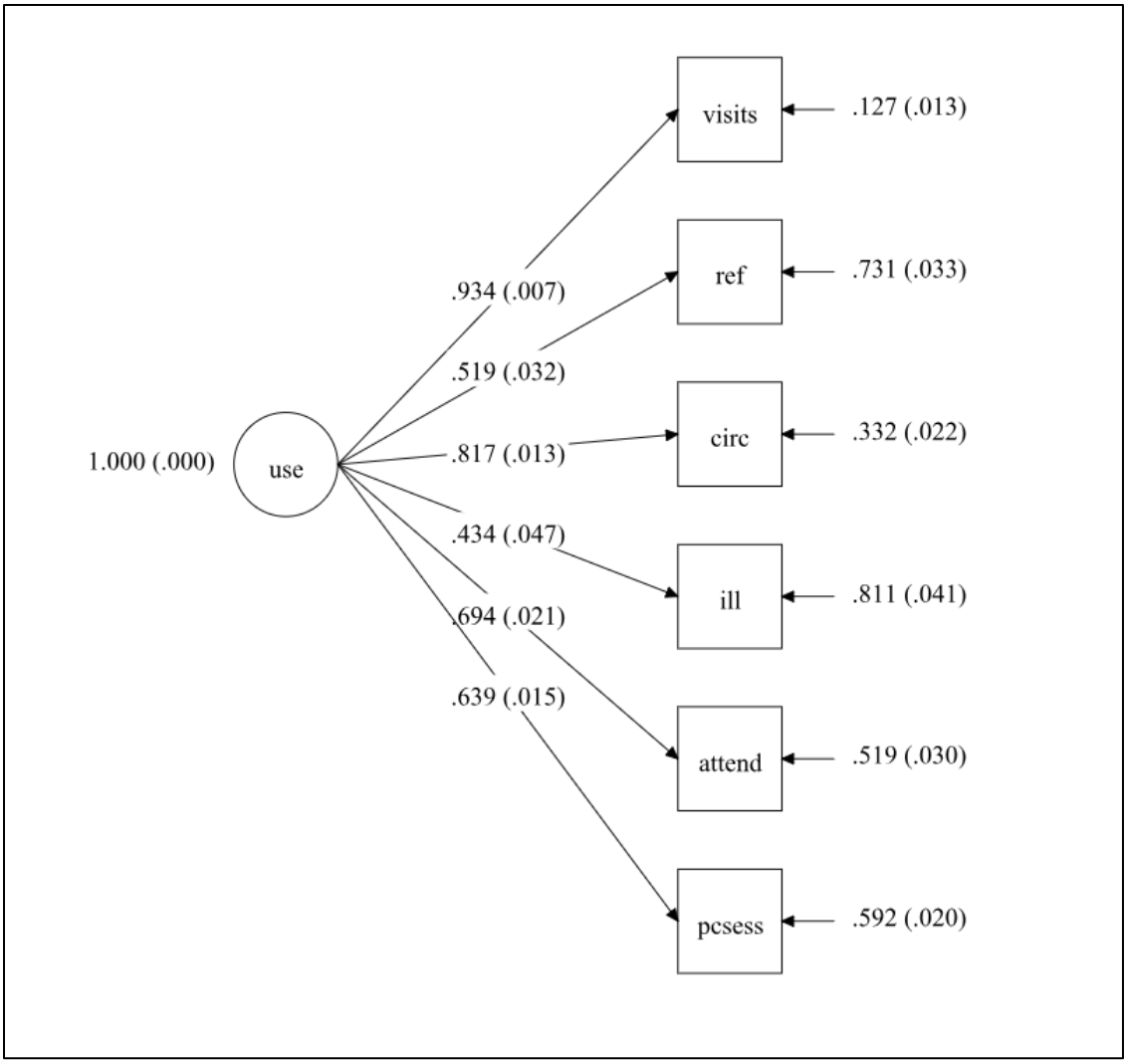

Figure 11. Standardized coefficients for a six-indicator model measuring Public Library Use. Standard errors for parameter estimates are shown in parentheses.

As with the Service Availability factor, all measurements of model fit and reliability suggest that including the public computer variable included in the PLS data set reduces the fit for a single-factor measure of public library use. When compared with the measurement of Service Availability, the measures of library use display better model fit, but generally lower factor loadings and measures of construct reliability/variance explained. Also in contrast to the Service Availability factor, the additional variable

added in the six-factor model is not the lowest-loading indicator in either model. The indicator with the lowest loadings is instead the inter-library loan activity indicator, a variable that may be dependent on the strength of regional lending services, rather than library-specific service availability.

Correlations between factor scores and observed predictors. Factor scores for the five-indicator Use factor and the five-indicator Service Availability factor were saved and correlations between these factors as well as other hypothesized observed predictors of use were obtained using a combined measurement model estimating no directional relationships other than factor loadings. These correlations are presented in Table 9.

			1	2	3	4	5	6	7	8
1	ln of Operating Revenue *per capita* (totrev)	Correlation	1	.971**	.419**	-.060**	.412**	-0.015	.856**	.930**
		Covariance	0.506	0.481	0.711	-0.003	0.044	-0.002	0.286	0.413
2	ln of operating expenditures *per capita* (provided for reference)	Correlation	.971**	1	.436**	-.072**	.436**	-0.011	.881**	.955**
		Covariance	0.481	0.485	0.724	-0.004	0.045	-0.001	0.288	0.415
3	Local median income (income)	Correlation	.419**	.436**	1	.262**	.768**	-.178**	.403**	.426**
		Covariance	0.711	0.724	5.679	0.049	0.273	-0.06	0.451	0.634
4	Percentage of local households with children (child)	Correlation	-.060**	-.072**	.262**	1	-.072**	-.218**	-.113**	-.095**
		Covariance	-0.003	-0.004	0.049	0.006	-0.001	-0.002	-0.004	-0.005
5	Percentage of local population with a college education (edu)	Correlation	.412**	.436**	.768**	-.072**	1	-.127**	.460**	.449**
		Covariance	0.044	0.045	0.273	-0.001	0.022	-0.003	0.032	0.042
6	Percentage of local population identifying as white (race)	Correlation	-0.015	-0.011	-.178**	-.218**	-.127**	1	.121**	.073**
		Covariance	-0.002	-0.001	-0.06	-0.002	-0.003	0.02	0.008	0.006
7	Factor scores for Use (use)	Correlation	.856**	.881**	.403**	-.113**	.460**	.121**	1	.965**
		Covariance	0.286	0.288	0.451	-0.004	0.032	0.008	0.221	0.283
8	Factor scores for Service Availability (serv)	Correlation	.930**	.955**	.426**	-.095**	.449**	.073**	.965**	1
		Covariance	0.413	0.415	0.634	-0.005	0.042	0.006	0.283	0.389

*Table 9. Correlations and covariances between saved factor scores and proposed predictors. N= 5,144. Pearson correlations. ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).*

Significant and strong correlations were found between the total Revenue variable, the Use factor, and the Service Availability factor. These correlations support the assertion that the Revenue-Use relationship may be mediated by the Service Availability factor. Other notable observations captured in the above table include the weak but negative correlation between the Use factor and ratio of households with children in the matched local area, the difference in the strength of the relationship between the race variable and Use versus the race variable with the Services factor, and the race variable with the revenue predictor. Income is strongly and positively correlated with the percentage of the local population with a college education. The measure of median income is also moderately and positively correlated with operating revenue, expenditures, and the factor scores for Use and Service Availability. The income variable is also weakly but significantly negatively correlated with the percentage of the local population identifying as white, which may indicate that higher median incomes are associated with more racially diverse areas or geographies.

A note on interpreting model estimates.

The parameters for the models tested in the present research have been presented as standardized coefficients. However, if readers are interested in interpreting the models in real terms, they should bear in mind the implication of data transformation. As noted in Chapter 3, all indicators of the latent factors and the Revenue variable were transformed to their natural logarithms to reduce the effect of outlying cases on the overall distributions being compared. A loglinear transformation creates a new score which is the power to which the natural number *e* (also known as Euler's number) must be raised to equal the raw number. This reduces the effect of outlying cases while preserving rank

order. Number *e* is a non-rational number close to 2.72 that is the base for equations modeling constant exponential growth. While public service organizations such as libraries do not experience constant growth, their size is closely related to the size of populations whose growth can be modeled using these equations.

In practice, this means that the marginal effect of each additional dollar per-capita of operating revenue is very large for low values, but that values greater than the natural number *e* demonstrate diminishing returns with regard to the levels of Service Availability and Use that they can "buy." Likewise, the Service Availability and Use factors predict indicators that follow a similar distribution: a factor score close to *e* for use would predict indicator counts that correspond with the unstandardized factor loadings while larger changes in factor scores will have a multiplicative effect on the exponentiated factor loadings and subsequent estimates of the factor indicators. To demonstrate this relationship a logarithmic distribution for raw values between the asymptote of zero and 10 is presented in Figure 12, with dashed lines indicating the intersection of x=*e* (approximated as 2.72) and y=1.

The University of California, Los Angeles (UCLA) Statistical Consulting Group has published an in-depth discussion of the interpretation of regression weights for logarithmically transformed variables at https://stats.idre.ucla.edu/other/mult-pkg/faq/general/faqhow-do-i-interpret-a-regression-model-when-some-variables-are-log-transformed/ (UCLA Statistical Consulting Group, 2018).

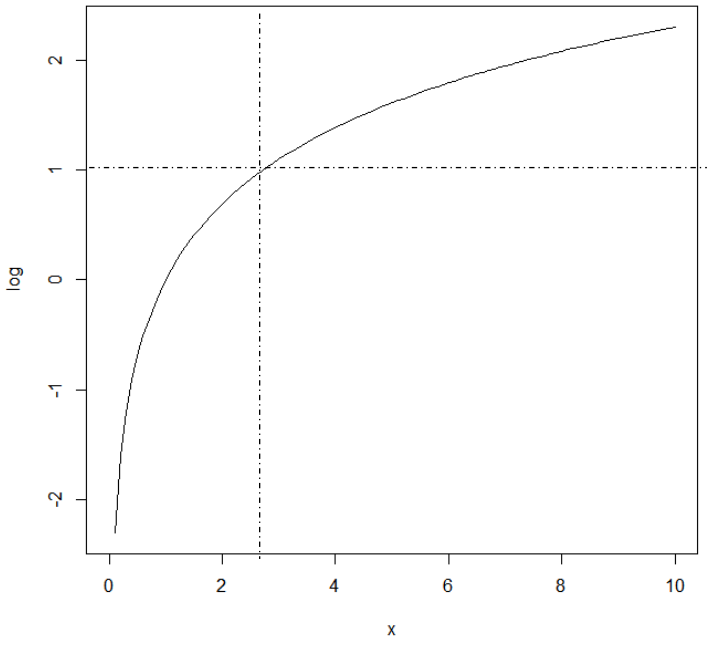

Figure 12. Plot of the natural logarithmic transformation.

Structural equation models. The two structural equation models tested are presented along with standardized parameter estimates in Figures 13 and 14 below. Measures of fit (including measures for the baseline model, which assumes all regression weights and disturbances to be 0 and all factor loadings to be 1) are presented in Table 10. To evaluate the difference in fit between these two models, a difference test was conducted evaluating the change in chi-square statistics between the partial mediation model and the nested model assuming complete mediation of the revenue-use relationship through the Service Availability variable. The result of this test was significant, with a Satorra-Bentler Scaled Chi-square Difference (df=1) of 389.45, $p<.001$. This implies that although there is strong evidence that the Service Availability factor mediates the relationship between Revenue and Use, the structural model implying full mediation displays a statistically significant worse fit than the partial mediation model, even if most measures of model fit imply an acceptably-fitting model.

	Baseline Model			Model Assuming Partial Mediation			Model Assuming Full Mediation		
	value	df	p-value	value	df	p-value	value	df	p-value
Chi-square	13759.34	95	<.0001	1557.658	78	<.0001	1655.069	79	<.0001
RMSEA	0.167			0.061			0.062		
SRMR	0.378			0.062			0.066		
TLI	0			0.868			0.861		
AIC	65231.11			23692.92			24083.31		
Scaling correction factor	3.571			4.8546			4.8059		

Table 10. Summaries of model fit for baseline structural model, partial mediation, and full mediation model.

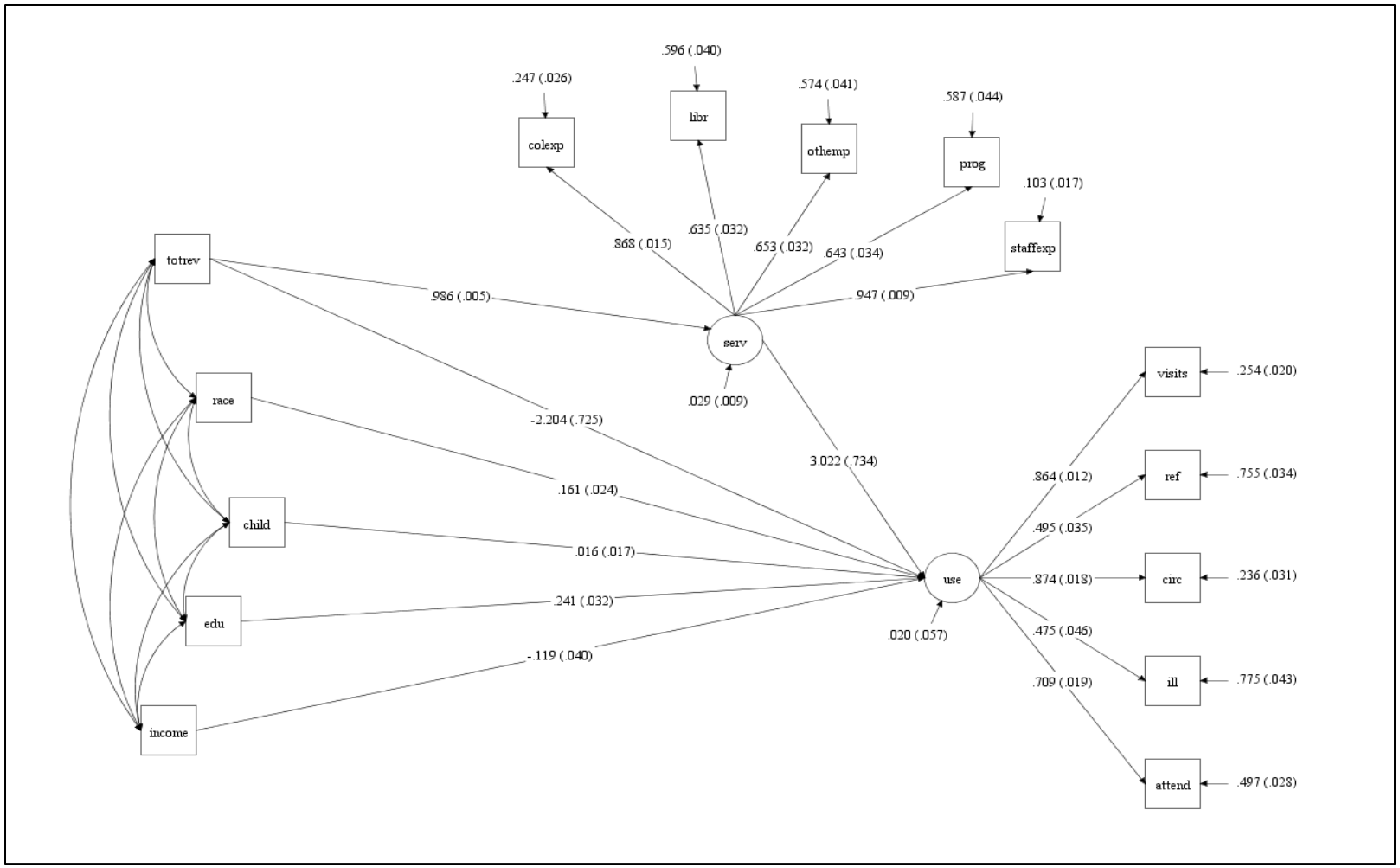

Figure 13. Standardized coefficients for a structural equation model of library action within an environmental context assuming partial mediation of the relationship between Operating Revenue and Public Library Use by Service Availability.

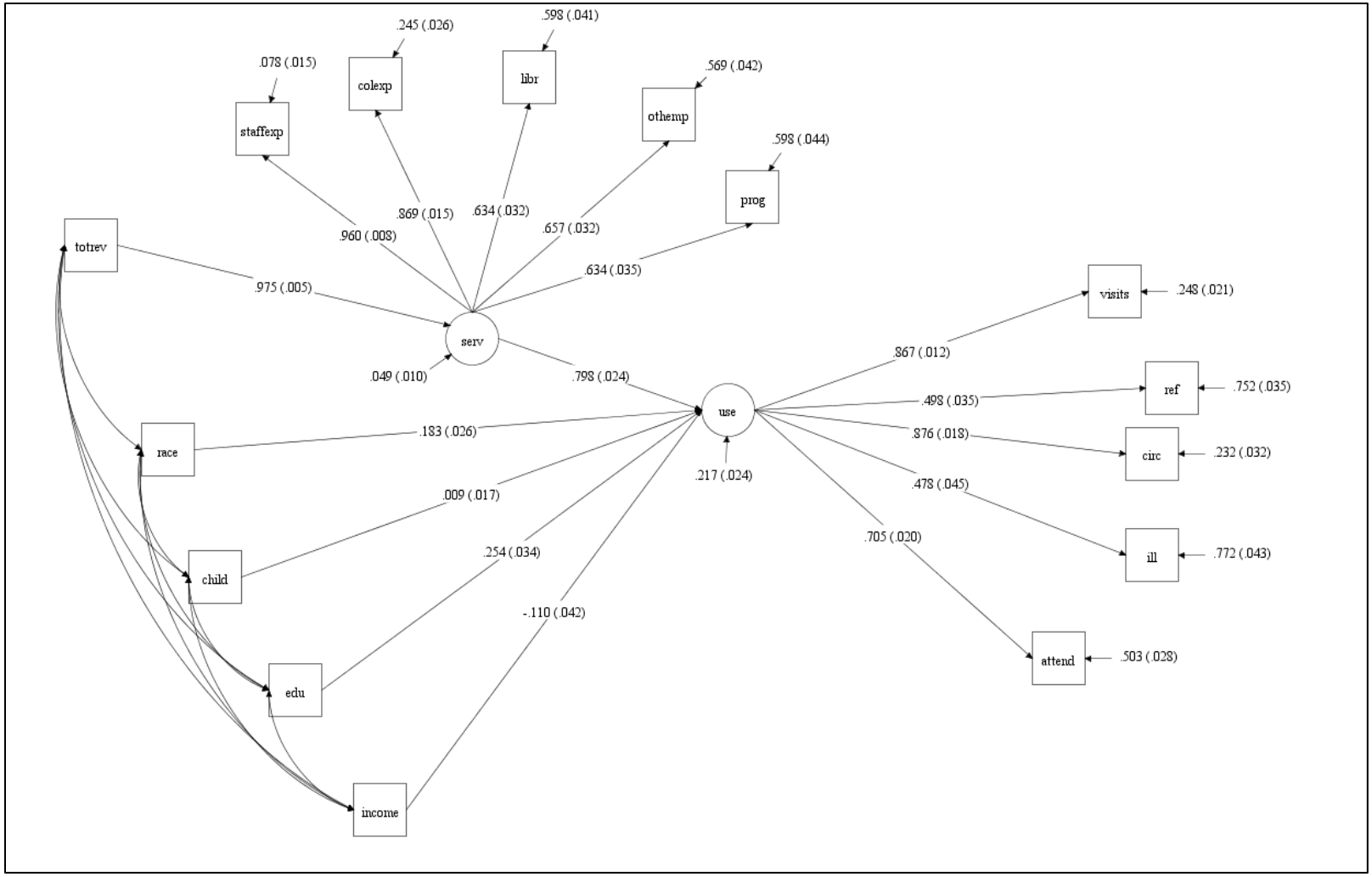

Figure 14. Standardized coefficients for a structural equation model of library action within an environmental context assuming full mediation of the relationship between Operating Revenue and Public Library Use by Service Availability.

Chapter 5: Discussion

Responses to Planned Research Questions

Is a factor structure measuring public library service availability supported by the PLS data set? In Chapter Four, two factors predicting several measures of public library use were proposed and examined for evidence of fit and reliability. Both models fit the data moderately well and explained a meaningful portion of their total variance. This evidence could be taken as support for the construct validity of the factor structures proposed. The evidence presented above shows that the indicators included in the models are positively correlated with each other and – at least in aggregate at the national level – load onto a single predictor. This, in turn, suggests that service availability at public libraries could be discussed as a single concept. However, the procedures described above do not rule out the possibility of other measures that might better capture the full range of services available at public libraries or provide more practical utility when incorporated into a structural model linking monetary inputs to library use outputs.

As noted above, indicators available in the PLS data set that are arguably related to the construct include several alternative methods of counting employees, including the number of employees with graduate degrees, counts of materials held in library collections, and the number of programs targeted specifically at children or young adult audiences. While these variables could be useful for other research questions, such as the value of a graduate degree to library organizations or the relative effect of children's programs on library visitation compared to that of other programs, these variables were excluded from the models examined in the present report. The reason for this was an assumption that the variables did not increase the number of use dimensions being addressed but rather represented either narrowly-defined counts of aspects already covered by other counts (as in the case of counts of masters-degree-holders and children's programs) or difficult-to-interpret indicators (counts of physical materials might represent high levels of access or poor collection upkeep practices; counts of electronic collections are hard to translate into an amounts of consumable media without more information than is in the PLS). Limiting the number of variables for use in the model similarly limited the number of identifiable structures for analysis. With five indicators, only a first-order factor structure (in which all indicators load on a single factor) is possible unless some indicators load on more than one factor. The number of identifiable parameters in a model is equal to n*n+1, where n represents the number of variance/covariance relationships between observed. Thus, adding more indicators has a multiplicative effect on the number of structures that might be analyzed. Testing all possible factor models for an expanded set of candidate use indicators would be time consuming and potentially unproductive. However, separation of monetary and non-

monetary indicators of use could lead to factor structures with more easily-interpretable loadings. A more nuanced breakdown of expenditures could provide researchers with more information about which expenditure categories covary; a factor model for service availability that measures only non-monetary indicators might provide more additional information in an action model that includes monetary inputs.

Is a factor structure measuring public library use supported by the PLS data set? As with the Service Availability factor, the two models measuring public library Use demonstrated reasonable levels of fit, acceptable factor loadings, and measures of reliability. Fit measures indicate that the Use factors proposed fit the data in the filtered data set very well. However, factor loadings and measures of reliability suggest that alternate models might better explain the variance in the hypothetical Use factor, with some loadings estimated as lower than .5 and composite reliability/average variance extracted statistics indicating that the indicators in the model only explain around 50 percent of the variance in the latent factor. While the six-indicator model does have slightly worse fit and reliability statistics, the indicator added to the model (PC sessions per capita) did not have the lowest factor loading; the inter-library loan activity indicator was the least inter-correlated variable. This pattern might be attributable to an unanalyzed relationship between interlibrary loan frequency and the strength of regional lending services rather than library-specific service availability. Not only might further attempts to identify a Use measurement model benefit from testing alternative factor structures. However, developing a systematic process for selecting model component indicators could also benefit researchers by confronting assumptions about data collection consistency and indicators' logical connection to the units under analysis. As with the

Service Availability variable, the Use models proposed demonstrate reasonable fit and reliability, but neither should be considered an ideal quantitative model of library use; alternate factor structures could likely be fit, particularly if more indicators can be reasonably included. A single factor with more indicators, while likely being less well-fitting could provide useful information about the degree to which different types of library services are used while a multi-factor Use model could allow use of library facilities (programs, PC use, etc.) to be considered separately from media services and research help services.

Does service availability fully mediate the relationship between public library revenue and public library use? The two latent factors incorporated into the structural model both had strong positive correlations with the revenue predictor (see Table 3). When incorporated into a model assuming partial mediation of the Revenue → Use relationship by the Service Availability variable, an inconsistent mediation effect was revealed in which the direct effect between Revenue and Use was measured to be negative, while the Revenue-Service Availability and Service Availability-Use relationships were measured to be positive. Practical interpretation of inconsistent mediation effects can be difficult. MacKinnon, Fairchild, and Fritz (2007) note that inconsistent mediation patterns are, "more common in multiple mediator models where mediated effects have different signs." However, in the case of the Revenue, Service Availability, and Use relationship, all bivariate effects indicate positive associations.

According to Kenny's (2018) discussion of structural mediation, the total effect in a partial mediation model can be described as the sum of the direct mediation effect and the product of the indirect effects:

$$\mathbf{c = c' + ab}$$

where **c** represents the total effect, **c'** represents the direct effect of the causal variable on the outcome variable, **a** represents the effect of the causal variable on the mediating variable, while **b** represents the path between the mediating variable and the outcome variable.

When the coefficients estimated by the partial mediation structural model are used in this equation, a total effect of β =.776 is estimated for the mediated relationship. This hews very closely to the estimate of the total effect seen in the full mediation model (where the direct effect **c** would be assumed to be zero) of β =.778. The direct, indirect, and total estimated effects for the tested models are shown in Table 11.

	Model Assuming Partial Mediation	Model Assuming Full Mediation
Revenue-Use (direct effect)	-2.204	-
Revenue-Service Availability-Use (indirect effect)	3.022*.986 = 2.98	.975*.798 = .778
Total Estimated Effect	0.776	0.778

Table 11. Direct, indirect, and total effects of operating revenue on the Use factor in models assuming partial and full mediation.

There appears to be strong evidence that the Service Availability fully mediates the revenue-to-use relationship, and that the large negative coefficient seen in the model assuming partial mediation might best be interpreted as acknowledging that there is a weaker correlation between Service Availability and Use than between Revenue and Service Availability, as seen in the model assuming full mediation. Alternatively, the negative direct effect could be interpreted as an indication of the importance of library services to library operations. While the effect of operating revenue on services is very strong, the effect of services on use is stronger. The inconsistent sign observed in the direct effect of revenue on use could be interpreted as an indication that revenue does not always translate to use; the services provided by the library are a crucial component of the library use process.

The simpler "full mediation" model presents the same relationships but from a different perspective. This model has statistically significant worse fit based on adjusted chi-square differences. However, its simplicity makes the process of library service provision and associated use patterns easier to communicate, if the goal is to emphasize predictors of use other than library service availability. The line of effect from revenue to services to use is still strong, but the magnitude of the path between services and use is not so large as to overwhelm the contributions of community effects on use in visual depictions. While the models present mathematically similar stories, they emphasize different aspects of the library organization action model.

What predicts public library use? When considering all predictors of use in the full models (both the full mediation model and the partial mediation model), the local descriptive variables taken from matched ACS model do not predict use as strongly as the Service Availability factor or the co-linear Revenue predictor. However, moderately strong associations were found between Use and three of the four local descriptive variables: the percentage of persons identifying as white in the local census-designated place or county, the local median income, and percentage of persons with college educations in the local areas were all positive predictors of library use, while the percentage of local households with children was not predictive of changes in the Use factor.

Operating revenue and service availability as predictors of library use. As noted in the previous subsection, operating revenue is mediated by the service availability factor. Both predictors have strong positive associations with the estimated Use factor, but revenue is not universally directed toward services that are used. Given stable revenue, a smaller difference between the path between revenue and Service Availability compared to the path between Service Availability and Use could indicate improvements in library management practice and/or the adoption of library services by local communities.

Local median income as predictor of library use. Despite being having a strong positive correlation with operating revenues and having a moderate positive correlation with the Use factor score derived from the measurement model, local median income had a small negative association with the Use factor in the structural model. This may indicate that libraries located in wealthier areas see less demand for their services than those located in lower income communities, despite being able to afford more services. However, it should be noted that most public libraries in the United States derive a large portion of their funding through local property taxes rather than income taxes and that local tax rates and property values (as opposed to net local revenue) are not included in the PLS data set. Gini coefficients or other measures of income distribution were also not included in the models tested. While the negative relationship between local median income and use is noteworthy, a more focused investigation using detailed data could reveal more about the nature of this relationship.

Percentage of local households with children as predictor of library use. The relationship between the percentage of local households with children and the Use factor was positive, but very small. This highlights the difference between the phenomenon of individual library use and local library use; while children's programs are known to be a major driver of library attendance (Pew Research Center, 2016), the percentage of households with children in a given locality is not highly variable within the continental United States. What variation that does exist does not appear to be associated with use of local libraries. Despite the limited magnitude of the relationship, it is worth noting that the percentage of households with children had a small negative relationship with revenue and the Service Availability factor, as well as the Use factor when not controlling for revenue. While the reason for this disparity cannot be determined in the present study, it merits further study.

Percentage of local population with a college degree as predictor of library use. The relationship between the percentage of the local population with a college degree – a rough proxy for the education level of the local population – and the Use factor was positive and significant. This variable is strongly correlated with local median income, which complicates its interpretation. However, the difference in the direction of these variables' relationship with the Use factor suggests that community engagement with formal education may also translate to engagement with public libraries.

Percentage of local population identifying as white as predictor of library use. Examination of the table of bivariate correlations (Table 3) between the predictors and other factors in the models reveals that while education and income have moderately strong correlations with revenue and the Service Availability factor, the percentage of the local population identifying as white is only significantly correlated with the Use factor. This may indicate disparities in access to library services in some communities with large minority presence, a mismatch between library service offerings and community demands, or another cause rooted in a library area descriptor linked to racial demographics not included in the models tested, such as the size of the library service area or the urban density of the locality served by the library administrative entity.

Interpretation of community and organization data versus individual data. I believe that should be emphasized again that the data underlying the models described above describes large areas: counties, cities, and library service areas that contain many individuals. While community outreach on the part of a library organization may play a role in an individual's decision to visit a library, the individual decision to use the library is not captured in any of the models described above. Instead, the model describes the process of converting revenue into services within the context of several local descriptors. This is useful for identifying describing relationships at the national level while revealing relationships in need of closer examination such as the disparity between revenue and community demand in the case of local households with children, or potential limits to access facing communities with large non-white populations.

Applying Findings from this Study

Theory-building. Many implementations of confirmatory factor analysis are used to provide supporting evidence for a proposed theory. While evidence supporting a model of public library action is presented in this report, the research described above might be better described as an investigatory process than an attempt to confirm or falsify the model proposed. The finding that library operating revenue predicts library service availability and use of services should not be surprising; understanding the degree to which these factors are correlated provides a baseline from which to build further theory.

The process of fitting available data to a theoretical framework prompts reconsideration of the logical relationship between observed measures of library services and use and the reality of library service provision. This in turn prompts questions that might lead to more robust models and data collection practices in future work. In the case of the present data, the patterns that can be seen in the models analyzed raise questions about how measurement of service capacity could be accomplished using non-monetary measures, whether it might be possible to dis-entangle local descriptors that predict library use from those that predict library revenue, and whether a multi-factor model for public library use might be more reliable than the single-factor models proposed. Although the models tested fall within acceptable guidelines with regard to model fit, further iteration of data point selection and model design in response to the process of interpreting the initial models could prove beneficial in terms of producing a more cohesive model for measuring library processes.

Applying findings in practical contexts. Although the findings presented above do not – in the author's opinion – represent a fully realized theory of library operations, the findings from the initial models presented above nevertheless offers several insights. These include the strong covariance of staff and collection expenditures, the identification of an apparent relationship between library use and local racial demographics, the quantification of direct and indirect effects of revenue on use when mediated by a service availability factor. These findings could be used as the basis for further research or as points to support arguments for increases to funding or services.

The models tested in the present study also provide a solid foundation for further development of formal library data collection procedures, as model fit results indicate the potential for reliable multivariate measures of library use and service availability. Further development of these measurements could be used in conjunction with library valuation studies, assessments of current library services in relation to local literacy, information, and media needs, taxpayer return-on-investment studies, and assessments of library impacts.

Limitations of the Present Study

Although several issues with data selection, matching, and interpretation have been noted in previous sections, the limitations of this study should be addressed in a separate section so as to provide a reference point for future research. These limitations can be grouped into three main categories: data preparation, variable selection, and generalizability.

The data preparation process involved several filters and transformation processes that excluded library organizations with very large and very small service areas, as well

as libraries that did not report variables included in the model. For these reasons, libraries with very small or volunteer-heavy staffs were excluded from the analysis data set, as were library administrative units that cover very large populations. Data missingness was assumed to be primarily dependent on library organization size (a factor not included in the models tested), but the degree to which data were missing at random from the data set prior to filtering for data availability was not assessed. Libraries in remote locations and libraries in Ohio were also excluded, leading to an analysis data set that might not be representative of nationwide aggregates.

Selection of variables for the models presented in this report was undertaken with the goal of balancing representation of the various aspects of use and services captured by PLS data with model simplicity. As noted above, increasing the number of factor indicators could increase the number of potential models that could be compared. Alternative models for Service Availability and Use not assessed in the present report might provide better fit and/or more actionable information.

Another limitation worth noting is that the models presented show aggregate results but lack comparison groups differentiated by region, administration style, local population density, or time period. This limits the degree to which claims of generalizability of the models and their estimates might be made in the present report.

Areas for Further Research

There are several directions that further research related to this report might take. Most of the limitations listed above could be addressed if not resolved through reexamination of the 2015 PLS data. Alternative data preparation methods could be used, alternative factor models proposed, and tests of model invariance between sub-sections of

the library population could be undertaken. Pursuing this direction of research could help to refine the research community's understanding of the data set while providing librarians and library researchers with a richer understanding of library services and use patterns.

Other directions further research might take include more intensive examination of the role physical and social geographies play in determining use patterns. Combining a structural use model with a geographic analysis component could also lead to a more precise understanding of the relationships between library location, neighborhood demographics, library services, and library use, building on work such as that presented by Koontz et al. (2009). A related course of research could examine the effect of different funding models on library administrative entities' ability to provide services to their communities; libraries funded by local taxes at the township level might exhibit higher levels of service disparity than libraries funded through county or regional sources.

Finally, incorporation of a reliable and valid use factor in a longitudinal model that combines several years of PLS data could yield additional information about nationwide trends in library use over time. This effort might be complicated by changing data definitions over time, as well as by changing library use patterns and might work best incorporated into a multi-stage research project.

Conclusion

The study described in this report demonstrates a method for understanding publicly available administrative data through confirmatory factor analysis and structural equation modeling. The models proposed demonstrate reasonable levels of fit and

reliability, yielding actionable estimates for the relationships between library use, service availability, operating revenue, and select demographic descriptors. The models tested show a strong but not perfect predictive relationship between revenue and use when mediated by a latent factor representing services while identifying an apparent negative relationship between the Use factor and the percentage of non-white persons in the local community, which could point toward further research on equitable access to library services. However, I believe that reflecting on the process of examining the data and framing it within the context of an action model might lead to more impactful research on public data and public library use. At several points in this chapter, I have suggested that alternative models for the measurement of Service Availability and Use might be better-fitting or more explanatory than those presented in the report. I have also noted that further analysis of the consistency of the model's estimates between sub-sections of the population of library administrative entities represented in the Public Libraries Survey is merited before making any strong claims about the validity of the model and its estimates. The process of exploring the data within the context of the structural model highlighted these issues; an algorithmically-driven exploratory process would be oriented toward identifying the best linear combinations of model variables, while questions about the practical value of including or excluding a given indicator might go unaddressed. It was the process of fitting available data to a theoretical framework that led me to reconsider the logical relationship between observed measures of library services and use and the lived reality of library service provision.

A well-known aphorism in quantitative modeling is that, "all models are wrong; some models are useful." This phrasing is commonly attributed to George Box, who employed it to highlight the problem of overparameterization of models (Box & Hunter, 1978). I feel that it is appropriate to cite Box's words in this discussion. In my opinion, the goal of quantitative modeling is not to perfectly describe the observed world but to describe it in such a way that useful connections can be made, either by the researcher or someone else. While the initial parameter estimates reported here may be used to provide a baseline for comparisons between library administration entities or as a way to re-frame the conversation around taxpayer return-on-investment, I believe that its greatest use will be realized as a first step in the generation of more useful models and research approaches.

References

Aabø, S. (2005a). Are public libraries worth their price? A contingent valuation study of Norwegian public libraries. New Library World 106 pp 487-495.

Aabø, S. (2005b). Valuing the benefits of public libraries. Information Economics and Policy 17. pp.175-198. doi:10.1016/j.infoecopol.2004.05.003

Adamowicz, W., Louviere, J., & Williams, M. (1994). Combining revealed and stated preference methods for valuing environmental amenities. *Journal of environmental economics and management*, *26*(3), 271-292.

Adedokun, O. A., Childress, A. L., & Burgess, W. D. (2011). Testing conceptual frameworks of nonexperimental program evaluation designs using structural equation modeling. American Journal of Evaluation, 32(4), 480-493.

Alkin, M.C. (Ed.) (2013). Evaluation Roots Thousand Oaks, CA: Sage.

American Academy of Arts and Sciences. (2016). Use of Public Libraries. Retrieved from https://www.humanitiesindicators.org/content/indicatordoc.aspx?i=98

American Academy of Arts and Sciences. (2016). Public Library Revenue, Expenditures, and Funding Sources. Retrieved from https://www.humanitiesindicators.org/content/indicatordoc.aspx?i=419

Arns, J. W., Williams, R., & Miller, K. (2013). Assessing the Value of Public Library Services: A Review of the Literature and Meta-Analysis (META). *School of Library and Information Science, University of South Carolina, Columbia, SC.* Retrieved from http://www.libsci.sc.edu/metaweb/ValuingPublicLibraries_FinalReport.pdf

Barron, D.D., Williams, R.V., Bajally, S., Arns, J., Wilson, S. (2005). The Economic Impact of Libraries in South Carolina. Retrieved from http://www.libsci.sc.edu/sceis/final%20report%2026%20january.pdf

Bhatt, R. (2010). The impact of public library use on reading, television, and academic outcomes. *Journal of Urban Economics 68*. pp. 148-166. doi:10.1016/j.jue.2010.03.008

Box, G. E. P. Hunter, W. G. (1978), *Statistics for Experimenters*. Hoboken, NJ: John Wiley & Sons.

Brewster, L. (2014). The public library as therapeutic landscape: A qualitative case study. *Health and Place 26* pp. 94-99.

Buckland, M. K. (1999). Library services in theory and context. London: Elsevier. Retrieved from http://digitalassets.lib.berkeley.edu/sunsite/Library%20Services%20in%20Theory%20and%20Context,%202nd%20Edition.pdf

Cameron, T. A., Poe, G. L., Ethier, R. G., & Schulze, W. D. (2002). Alternative non-market value-elicitation methods: are the underlying preferences the same? *Journal of Environmental Economics and Management*, *44*(3), 391-425.

Carson, R. T., Flores, N. E., & Meade, N. F. (2001). Contingent valuation: controversies and evidence. *Environmental and resource economics*, *19*(2), 173-210.

CDC Program Performance and Evaluation Office. (2016). Overview of the Office. Retrieved from https://www.cdc.gov/program/overview/index.htm

Conniff, R. (2012). What's wrong with putting a price on nature? *Yale Environment 360* Retrieved from http://e360.yale.edu/features/ecosystem_services_whats_wrong_with_putting_a_price_on_nature

Chen, H. T. (1990). *Theory-driven evaluations*. Sage.

Davies, T. (2010). *Open Data, Democracy and Public Sector Reform: A Look at Open Government Data Use from data.gov.uk.* Retrieved from http://www.opendataimpacts.net/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf

De Blaeij, A., Florax, R. J., Rietveld, P., & Verhoef, E. (2003). The value of statistical life in road safety: a meta-analysis. *Accident Analysis & Prevention*, *35*(6), 973-986.

De Witte, K., & Geys, B. (2013). Citizen coproduction and efficient good provision: theory and evidence from local public libraries. *European Journal of Operational Research 224* 592-602.

Diamond, P.A. & Hausman, J.A. (1994). Contingent valuation: Is some number better than no number? *Journal of Economic Perspectives 8*(4) pp. 45-64.

Dworkin, (1980). Is wealth a value? *The Journal of Legal Studies 9*(2). pp. 191-226. Retrieved from http://www.jstor.org/stable/724129

Dyer, N. G., Hanges, P. J., & Hall, R. J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *The Leadership Quarterly*, *16*(1), 149-167. doi:10.5465/AMBPP.1997.4989260

Edwards, J.R., & Lambert, L.S. (2007). Methods for integrating moderation and mediation: A general analytic framework using moderated path analysis. *Psychological Methods 12,* 1-22.

Evans, G.E., Amodeo, A.J., & Carter, T.L. (1999). *Introduction to Library Public Services* (6th ed.). Englewood, CO: Libraries Unlimited.

Evans, G.E., Saporano, M.Z., Christie, H., & Sinwell, C. (2015). *Library Programs and Services: The Fundamentals* (8th ed.). Denver, CO: Libraries Unlimited.

Farmer, L.S. & Safer, A.M. (2016). *Library Improvement through Data Analytics*. Chicago, IL: ALA-Neal-Schulman.

Fetterman, D.M. (2013). Empowerment evaluation: Learning to think like an evaluator. In *Evaluation Roots* (Ed: Alkin, M.C.). Thousand Oaks, CA: Sage.

Fischoff, B. & Davis, A.L. (2014). Communicating Scientific Uncertainty. *Proceedings of the National Academy of Sciences of the United States of America 111.* Retrieved from http://www.pnas.org/content/111/Supplement_4/13664.full

Fisher, K.E., Durrance, J.C., & Hinton, M.B. (2004). Information Grounds and Use of Needs-Based Services by Immigrants in Queens, New York: A context-based, outcome evaluation approach. *Journal of the American Society for Information Science and Technology 55*. pp. 754-766.

Fleeter, H. (2016). The Return on Investment of Ohio's Public Libraries. Retrieved from http://olc.org/wp-content/uploads/documents/post-id_2060/2016/04/Ohio-Public-Libraries-ROI-Report.pdf

Fogg, L. (2006). A Short History of Cost-Benefit and Cost-Effectiveness Analysis. Conference Presentation. Presented at the Chicago Chapter of the American Statistical Organization Cost Benefit and Cost Benefit Analysis Conference. May 5, 2006: Chicago, IL. Retrieved from http://community.amstat.org/zthechicagochapterold/calendar/pastevents/20052006/may52006conference/downloadpresentationshistoryofcostbenefitanalysis

Gaille, B. (2016). 10 Pros and Cons of Benchmarking. Retrieved from https://brandongaille.com/10-pros-and-cons-of-benchmarking/

Grimes, J., Manjarrez, C.A., Miller, K.A., Owens, T., & Swan, D.W. (2013). Public Libraries in the United States Survey Fiscal Year 2010. Retrieved from https://www.imls.gov/sites/default/files/publications/documents/PLS2010.pdf

Grimes, J., Manjarrez, C.A., Miller, K.A., Owens, T., & Swan, D.W. (2014a). Public Libraries in the United States Survey Fiscal Year 2011. Retrieved from https://www.imls.gov/sites/default/files/publications/documents/PLS2011.pdf

Grimes, J., Manjarrez, C.A., Miller, K.A., Owens, T., & Swan, D.W. (2014b). Public Libraries in the United States Survey Fiscal Year 2012. Retrieved from https://www.imls.gov/sites/default/files/legacy/assets/1/AssetManager/PLS_FY2012.pdf

Hallam, A.W., & Dalston, T.R. (2005). Managing Budgets and Finances: A How-to-Do-It Manual for Librarians and Information Professionals. New York, NY Neal-Schuman Publishers.

Hanemann, W.M. (1991). Willingness to Pay and Willingness to Accept: How much can they differ? The American Economic Review 81(3). pp. 635-647.

Harris, B.F. (2013). Taxpayer Return on Investment in Florida Public Libraries 2013. Retrieved from http://roi.info.florida.gov/Content/PDFs/Studies/Library%20ROI%202013.pdf

Hemmeter, J. A. (2006). Estimating public library efficiency using stochastic frontiers. Public Finance Review, 34(3), 328-348.

Holt, G. E., & Elliot, D. (2001). Public library benefits valuation study. Final report to the Institute of Museum and Library Services for National Leadership. St. Louis, MO: St. Louis Public Library.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural equation modeling: a multidisciplinary journal, 6(1), 1-55.

IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.

Imholz, S., & Arns, J. (2007). Worth their Weight: An Assessment of the Evolving Field of Library Valuation. New York, NY: Americans for Libraries Council.

Institute for Museum and Library Services (IMLS). (2016). Public Libraries in the United States Survey Fiscal Year 2013. Retrieved from https://www.imls.gov/sites/default/files/publications/documents/PLSfy2013.pdf

Institute for Museum and Library Services (IMLS). (2017). Public Libraries in the United States Survey Fiscal Year 2014. Retrieved from https://www.imls.gov/sites/default/files/publications/documents/plsfy2014.pdf

Ivy Group (2016). Who's Comparing? Benchmarking Library Performance. Retrieved from https://ivygroup.com/blog/benchmarking-library-performance/

Jenatabadi, H.M. & Ismail, N.A. (2014). Application of structural equation modelling for estimating airline performance. *Journal of Air Transport Management 40* pp. 25-33. http://dx.doi.org/10.1016/j.jairtraman.2014.05.005

Joo, S. & Cahill, M. (2017). The relationships between the expenditures and resources of public libraries and children's and young adults' use: An exploratory analysis of Institute of Museum and Library Services public library statistics data. *Journal of Librarianship and Information Science 2017.* https://doi.org/10.1177/0961000617709057

Kamer, P. M. (2005). Placing an economic value on the services of public libraries in Suffolk County, New York. Bellport, NY: Suffolk Cooperative Library System Retrieved from http://scls.suffolk.lib.ny.us/pdf/librarystudy.pdf

Kenny, D.A. (2015). Measuring Model Fit. Retrieved from http://davidakenny.net/cm/fit.htm

Kenny, D.A. (2018) Mediation. Retrieved from http://davidakenny.net/cm/mediate.htm

Kim, G. (2011). A critical review of valuation studies to identify frameworks in library services. Library and Information Science Research 33. pp. 112-119. doi:10.1016/j.lisr.2010.09.006

Kline, R.B. (2011). *Principles and Practice of Structural Equation Modelling* (3rd ed.) New York, NY: Guilford Press.

Koontz, C. M., Jue, D. K., & Bishop, B. W. (2009). Public library facility closure: An investigation of reasons for closure and effects on geographic market areas. *Library and Information Science Research*, *2*(31), 84-91.

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. Journal of the American society for information science, 42(5), 361-371.

Levin, H.M. (2013). Cost-effectiveness evaluation in education. In *Evaluation Roots*, Alkin, M. (ed.). Thousand Oaks, CA: Sage.

Levin, H.M. Driscoll, W. & Fleeter, H. (2006). Value for money: Southwestern Ohio's return on investment in public libraries. Retrieved from http://9libraries.info/docs/EconomicBenefitsStudy.pdf

Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological methods*, 10(1), 21.

Lyons, R., & Lance, K.C. (2014). Outputs, outcomes, and other data. Retrieved from http://lj.libraryjournal.com/2014/11/managing-libraries/lj-index/class-of-2014/the-star-libraries-2014

Institute for Library and Museum Services (IMLS). (2016). *PLS Web Portal User's Guide.* Retrieved from https://www.imls.gov/sites/default/files/pls_users_guide_fy2015_508.pdf

Institute for Library and Museum Services (IMLS). (2017). PLS (PLS) Data and Reports. Retrieved from https://www.imls.gov/research-evaluation/data-collection/public-libraries-survey/explore-PLS-data/PLS-data

McDonald, R. P., & Ho, M. H. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological methods*, *7*(1), 64.

Malliari, A., & Togia, A. (2017). An analysis of research strategies of articles published in Library Science journals: the example of Library and Information Science Research. *Qualitative and Quantitative Methods in Libraries*, *5*(4), 805-818.

Mehra, B., Black, K., Singh, V., Nolt, J., Williams, K. C., Simmons, S., & Renfro, N. (2017). The social justice framework in the information technology rural librarian master's scholarship program: Bridging the rural digital divides. *Qualitative and Quantitative Methods in Libraries*, 5-11.

Meijer, A.J., Curtin, D., & Hillebrandt, M. (2012). Open government: Connecting vision and voice. *International Review of Administrative Sciences 78*. DOI: 10.1177/0020852311429533

Microsoft Corp. Released 2016. Microsoft Excel 2016. Redmond, WA: Microsoft Corp.

Mitchell, R.C. & Carson, R.T. (1989). Using Surveys to Value Public Goods: The Contingent Valuation Method. Washington, DC: Resources for the Future.

Muthén, L. K. (2010). Forum post in response to: "Mplus Discussion: Multivariate normality". Retrieved from http://www.statmodel.com/discussion/messages/22/5308.html?1279319060

Muthén, L. K., & Muthén, B. O. (2012). Mplus. Statistical analysis with latent variables. Version 7. Los Angeles, CA: Muthén & Muthén

Nichols, J. (2005). PILOT payments: a potential revenue source for public libraries. The Bottom Line: Managing Library Finances 18. Doi: 10.1108/08880450510632235

Nashville Area Chamber of Commerce Research Center. (2016). Economic Impact and Contribution Analysis: Nashville Public Library. Retrieved from https://library.nashville.org/sites/default/files/field/page/file/nashville-public-library-impact-and-contribution_study.pdf

Office of Management and Budget (OMB) (2017). America First: A Budget Blueprint to Make America Great Again. Retrieved from http://www.budget.gov/

Organization for Economic Co-Operation and Development (OECD). (2006). Cost-benefit analysis and the environment: Recent developments. Retrieved from http://www.oecd.org/environment/tools-evaluation/36190261.pdf

Orr, R. H. (1973). Measuring the goodness of library services: A general framework for considering quantitative measures. Journal of Documentation, 29(3), pp. 315–332.

Owens, P.L., & Sieminski, M.L. (2007). *Local and State Sources of Funding for Public Libraries: The National Picture.* Williamsport, PA: RPA, Inc. Retrieved from http://www.pcblpa.org/Portals/8/Downloads/PaLALibraryFundingStudy.pdf

Pew Research Center. (2016). *Libraries 2016.* Retrieved from http://www.pewinternet.org/2016/09/09/libraries-2016/

Pooley, P.K., Harper, R., Neal,M., Lewis, R., Whitfield, J., & Scheibe, M. (2008). Taxpayer Return on Investment in Florida Public Libraries 2008. Retrieved from http://roi.info.florida.gov/Content/PDFs/Studies/Library%20ROI%202008.pdf

Rabiyath, S. (2010). General Concepts and Computation Methods of Efficiency. New Delhi: Centre for Economic Studies and Planning. Retrieved from http://ssrn.com/abstract=1701003

Radford, M.L. (2006). The Critical Incident Technique and the Qualitative Evaluation of the Connecting Libraries and Schools Project. *Library Trends 55*. Pp. 46-64.

Ranganathan, S.R. (1931). The Five Laws of Library Science. London, England: Edward Goldston. Retrieved from https://babel.hathitrust.org/cgi/pt?id=uc1.$b99721;view=1up;seq=13

Rasmussen Reports. (2014). Is the Internet Killing Public Libraries? Retrieved from http://www.rasmussenreports.com/public_content/lifestyle/general_lifestyle/august_2014/is_the_internet_killing_public_libraries

Research Institute for Public Libraries. (2017). RIPL Regional Colorado Conference Proceedings. July 31-August 1, 2017, Colorado Springs, CO. Resources available at: https://s.lrs.org/regional-resources.

Chen, H. T., & Rossi, P. H. (1980). The multi-goal, theory-driven approach to evaluation: A model linking basic and applied social science. *Social forces, 59*(1), 106-122.

Scitovsky, T. (1941). A note on welfare propositions in economics, in: Arrow, K.J., Scitovsky, T. (Eds.), 1969. *Readings in Welfare Economics*. Allen and Unwin, London, pp.390-401.

Shaping Outcomes (2006). Retrieved from http://www.shapingoutcomes.org/course

Sin, S.J., Kim, K. (2008). Use and Non-Use of Public Libraries in the Information Age: A Logistic Regression Analysis of Household Characteristics and Library Services Variables. *Library and Information Science Research 30*. 207-215.

Sin, S.J. (2009). Sin, S. C. J. (2009). *Structural and individual influences on information behavior: A national study of adolescents' use of public libraries* (Doctoral dissertation, University of Wisconsin--Madison).

Sin, S.J. (2011). Neighborhood Disparities in Access to Information Resource: Measuring and Mapping U.S. Public Libraries' Funding and Service Landscapes. *Library and Information Science Research 33* 41-53.

Sonnenwald, D. H., Wildemuth, B. M., & Harmon, G. L. (2001). A research method using the concept of information horizons. An example from a study of lower socio-economic students' information seeking behavior. The New Review of Information Behaviour, 2, 65-86.

Sputore, A., & Fitzgibbons, M. (2017). Assessing 'Goodness': A Review of Quality Frameworks for Australian Academic Libraries. *Journal of the Australian Library and Information Association 66*(3), 207-230.

Stapleton, L. M., McNeish, D. M., & Yang, J. S. (2016). Multilevel and single-level models for measured and latent variables when data are clustered. *Educational Psychologist*, *51*(3-4), 317-330.

Steffen, N., Lietzau, Z., Lance, K. C., Rybin, A., & Molliconi, C. (2009). Public Libraries--A Wise Investment: A Return on Investment Study of Colorado Public Libraries. Retrieved from https://www.lrs.org/documents/closer_look/roi.pdf

Streatfield, D., & Markless, S. (2011). Impact evaluation, advocacy and ethical research: some issues for national strategy development. *Library Review* 60(4), 312-327.

Stueart, R.D. & Moran, B.D. (2007). *Library and Information Center Management* (7th ed.) Westport, CT: Libraries Unlimited.

Talja, S., Keso, H., & Pietiläinen, T. (1999). The production of 'context' in information seeking research: a metatheoretical view. *Information Processing & Management*, *35*(6), 751-763.

Torrieri, N. (2014). ACS Design and Methodology Version 2.0 January 20, 2014. Retrieved from https://www2.census.gov/programs-surveys/acs/methodology/design_and_methodology/acs_design_methodology_report_2014.pdf

Torkington, N. (2010). *Rethinking Open Data: Lessons Learned from the Open Data Front Lines.* Retrieved from http://radar.oreilly.com/2010/02/rethinking-open-data.html

United States Geological Survey (2009). Economic Justification: Measuring Return on Investment (ROI) and Cost Benefit Analysis (CBA). Retrieved from https://www.fgdc.gov/initiatives/50states/newspbp/EconomicJustification_ROI-CBA-Tutorial_v2_052809_FinalVersa.pdf

United States Government Accountability Office. (2014). Agencies Included Key Elements of Cost-Benefit Analysis, but Explanations of Regulations' Significance Could Be More Transparent. Retrieved from https://www.gao.gov/products/GAO-14-714

UCLA Statistical Consulting Group (218). Faq: How do I interpret a regression model when some variables are log transformed? Retrieved from https://stats.idre.ucla.edu/other/mult-pkg/faq/general/faqhow-do-i-interpret-a-regression-model-when-some-variables-are-log-transformed/

Vakkari, P. & Serola, S. (2011). Perceived outcomes of public libraries. *Library and Information Science Research 34*. pp. 37-44. doi: 10.1016/j.lisr.2011.07.005

Vakkari, P., Aabø, S., Audunson, R., Huysmans, F., Kwon, N., Oomes, M., & Sin, S. C. J. (2016). Patterns of perceived public library outcomes in five countries. *Journal of Documentation, 72*(2). pp. 342-361.

Varheim, A. (2014). Trust in libraries and trust in people: Social capital creation at the public library. *Library Quarterly: Information, Community, Policy 84*. pp. 258-277.

Vitaliano, D. F. (1997). X-inefficiency in the public sector: the case of libraries. *Public Finance Review25* (6). 629-643.

Voelker-Morris, R. (2004). Outcome-based Evaluation: Practical and Theoretical Applications. Culurework 8 (4). Retrieved from http://pages.uoregon.edu/culturwk/culturework28.html

Wetzels, M., Odekerken-Schröder, G., & Van Oppen, C. (2009). Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. *MIS Quarterly*, 177-195.

Zickuhr, K., Rainie, L., Purcell, K., and Duggan, M. (2013). How Americans Value Public Libraries in Their Communities. Retrieved from http://libraries.pewinternet.org/2013/12/11/libraries-in-communities/

Zweizig, D., & Dervin, B. (1977). Public library use, users, uses: Advances in knowledge of the characteristics and needs of the adult clientele of American public libraries. *Advances in Librarianship 7*, 231-255.

Zweizig, D., Johnson, D. W., Robbins, J., & Besant, M. (1994). The Tell It Manual: The Complete Program for Evaluating Library Performance Chicago: IL: American Library Association.