Sin, S.-C. J., & Kim, K.-S. (2008). Use and non-use of public libraries in the information age: A logistic regression analysis of household characteristics and library services variables. *Library & Information Science Research*, 30(3), 207-215. doi: 10.1016/j.lisr.2007.11.008.

# Use and non-use of public libraries in the information age: A logistic regression analysis of household characteristics and library services variables

Sei-Ching Joanna Sin Kyung-Sun Kim

School of Library and Information Studies, University of Wisconsin-Madison, 600 North Park Street, Madison, WI 53706, USA

# Abstract

For public libraries to achieve effective strategic planning, they must know who uses the public library and who does not use it. This study examines the characteristics of users and non-users of the public libraries using socio-demographic data from the *Current Population Survey*, a nationally representative survey of over 50,000 households conducted during October 13-19, 2002, and library services data from the *Public Libraries Survey 2002*. The study finds 34 variables to be significant. These variables including factors that have not often been studied, such as distance from the library, age/school attendance status, use of other types of library, and public library expenditure per state capita. It is also worth noting that disadvantaged groups, including ethnic minorities, recent immigrants, and people with disabilities, were less likely to use public libraries. This was true even after other factors such as education and income were held constant. The study provides a national-level assessment of the under-served populations. It also offers triangulation to other existing research, particularly qualitative information behavior studies of specific groups.

# 1. Introduction

Since the spread of the modern library movement in the 19th century, the number of public libraries in the United States has grown into the thousands. By 2004, there were 9207 public libraries with 16,549 library service outlets (Chute et al., 2006). The libraries serve the information, education, and recreational needs of the public and the community (D'Elia, 1993). In an age when the inequality in wealth distribution is growing (Mishel et al., 2006; Zhu, 2007) and differentiated access to digital information persists (Fox, 2005; Fox & Livingston, 2007), the public library is an important institution for facilitating information access and bridging the digital divide (Bertot et al., 2006; Bill & Melinda Gates Foundation, 2004; Bishop et al., 1999; McCook, 1994).

#### 1.1. Problem statement

While the public library aims to facilitate information access to all, only a portion of the public uses its collections, computer facilities, reference services, or library programs. Understanding the identity of users and non-users of public libraries is critically important for the libraries' strategic planning. Findings on this topic, however, have been conflicting. Multivariate analyses at a national level are particularly rare, which limits the generalizability of most research findings. In light of this research gap, this study examines the characteristics of U.S. public library users and non-users in this information age. Specifically, it investigates whether the public library users and the non-users are significantly different in their characteristics, and if so, what are the direction and magnitude of relationships between each independent variable (i.e., household and library service variables) and the household use/non-use of public libraries.

This complements other information behavior studies, especially those with small samples and/or with a narrower focus on particular locale, by providing data from a nationally representative population. Studies of specific groups—such as ethnic minorities (Agada, 1999), the poor (Chatman, 1985), and urban youth (Agosto & Hughes-Hassell, 2005)—suggested the respondents often did not use public libraries. This study helps examine whether similar patterns in the use of public libraries exist across the country. In addition, this research has practical implications because it can be considered as a national environmental scan. States and local libraries often conduct environmental scans and community analyses in order to understand the users/non-users in their service areas (Cooper, Bolt, Lance, & Webster, 1993). Large-scale nationwide studies, however, have been rare. The current study addresses this gap and provides LIS professionals with an assessment of the user and non-user characteristics at a national scale.

# 2. Literature review

Since the *Public Library Inquiry*, one of the earliest large-scale library use/non-use investigations commissioned in 1946, research has repeatedly shown that only a portion of the public utilizes the public library (e.g., Berelson, 1949; Chen & Hernon, 1982). The *Public Library Inquiry*, for example, found that only about 18% of adults and fewer than 50% of children and young adults used the public library once a year. Using tabulated data, Berelson examined the characteristics of the library users. He analyzed variables such as age, education, sex, occupation, economic status, martial status, race, religion, residence, and opinion and community leadership. The study found that education is the most significant factor affecting the use of libraries. Library users were more likely to be young, educated, and middle class (Berelson, 1949).

In the 1970s, Zweizig brought increased sophistication to public library use/non-use research. Zweizig applied multivariate analyses to help control the influence of concomitant variables and introduced a composite measure to quantify library use. Since then, research in public library use/non-use has evolved in

various aspects. Improvements have been made in the areas of conceptual frameworks, measures of library use, and the variables tested. For instance, D'Elia (1980) developed and tested a hierarchical conceptual model that included factors such as awareness of the availability of library services, perception of the physical accessibility of these services, perception of the degree of difficulty involved in using these services, and the use of non-public libraries. Van House (1983) proposed a time allocation model of public library use. Other researchers have proposed various other measures of library use (D'Elia, 1980; D'Elia, 1981; Marchant, 1994).

Characteristics such as education level, income, age, and gender are often studied in library use/non-use research. With the increasing recognition of the complexity of human information behavior (e.g., Dervin & Nilan, 1986; Wilson, 1999), researchers have introduced other variables. These variables include information processing style, information needs and seeking, social network, media use, and their relationship to the library (Zweizig, 1973); educational plan (Kronus, 1973); perception of accessibility (D'Elia, 1980); childhood socialization (Powell, Taylor, & McMillen, 1984); childhood history (Lange, 1988); motivators (Marchant, 1994); and personality (Powell, 1984). Environmental factors such as distance from a public library and community size have also been examined (University of Michigan Survey Research Center, 1948). Nevertheless, most use/non-use studies focused on the user-side variables. Library services variables, often used in performance evaluation in LIS research, are rarely examined in the context of library use/non-use. Less is known about *if* and *how* the supply side of the issue (such as the number of library outlets or collection expenditure) influences citizens' decisions to use or not use the public library. There is clearly a need to incorporate the supply side of the equation into library use/non-use research.

While research on public library use/non-use has expanded in the past decades, the relationship between different variables and public library use/non-use still appears inconclusive. In addition to the dearth of study testing the supply-side variables, studies concerning user-side variables have generated conflicting findings. Other than the effect of education level, extant studies often disagree on the significance and direction of effect of most variables. For example, some studies found age to be significantly related to public library use—younger individuals are more likely to be library users. However, (Evans, 1970), (Kronus, 1973) and (Zweizig & Dervin, 1977) suggested that the effect of age is small and indirect. Sex was also found to be significant in predicting library use in D'Elia (1981), Marchant (1994), Powell et al. (1984), and Zweizig (1973), but not in Evans (1970), Kronus (1973), or Rees & Paisley (1968). Marital status was found to be significant in Kronus (1973) and in Marchant (1994), but not in Evans (1970), Powell et al. (1984), or Zweizig (1973). Socio-economic status such as income or occupation were often found to be significant (e.g. Westin & Finger, 1991), but others such as Kronus (1973) and Zweizig & Dervin (1977) indicated that this was probably due to the correlation between education and socio-economic status.

The differences in findings could be partly attributed to the differences in study design and data analysis methods (such as different measures of library use or the use of univariate analysis instead of multivariate analysis). Perhaps the most important difference, however, is the difference in sample population. The issue of external validity is one of the major obstacles in synthesizing the findings from previous research. Most library-use studies surveyed individuals in a particular geographical area. Marchant (1994), for example, cautioned readers against over-generalizing the result of his study. He rightly pinpointed the fact that information from one geographical area might not characterize the United States at large. A national-scale survey is clearly needed to strengthen the external validity of future studies.

There have been attempts to examine the characteristics of public library users/non-users by surveying a national sample (Campbell & Metzner, 1952; Gallup Organization, 1976; Gallup Organization, 1978; Lynch, 1997; Madden, 1979; Powell et al., 1984; Scheppke, 1994; Westin & Finger, 1991). However, with the exception of the multivariate analysis by Powell et al. (1984), the aforementioned national studies did not include inferential statistical analyses. Most of the national studies utilized the "variable-by-variable" approach, where the effect of different variables on library use is examined one at a time. As discussed in Zweizig (1973), the variable-by-variable approach might miss the correlation among user characteristics and might not be able to control for potential spurious effects. In-depth multivariate statistics have been successfully applied in some public library use research, as demonstrated by D'Elia (1980), Kronus (1973), Powell et al. (1984), and Zweizig (1973). To achieve a better understanding of the use/non-use of public libraries, more studies should build upon the theoretical and methodological achievement of extant in-depth studies. They should further expand them to examine a larger sample representative of the U.S. national population.

The rapidly changing information environment has also added to the need for an up-to-date library use/nonuse study. Abundant information is now made available through the Internet. This might lead to changes in public library use behaviors (D'Elia et al., 2002; Online Computer Library Center, 2004). In response, many public libraries are establishing an online presence and offering online services such as access to community information, electronic journals, digital collections, and virtual reference services through their websites (Bertot et al., 2006). As the information environment continues to evolve rapidly, there is a need to examine whether the user variables that were found to be significant in past studies are still significant in the present society. To better understand of the relationship between user characteristics and the use/non-use of the public library, more systematic investigations on a national scale are needed.

# 3. Procedures

This study employed logistic regression analysis to analyze the data drawn from the *Current Population Survey* October 2002: Survey on School Enrollment/Library Use Survey (hereafter referred to as CPS 2002) and the Public *Libraries Survey 2002.* The *Current Population Survey* is conducted monthly by the U.S. Census Bureau for the Bureau of Labor Statistics. *CPS 2002* was selected for this study. At the time of writing, it was the most recent *CPS* survey that included a topical survey on library use. The sample population of *CPS* consists of all persons in the civilian, non-institutional population of the United States who live in households. In the *CPS*, household is operationally defined as "all the persons who occupy a house, an apartment, or other group of rooms, or a room, which constitutes a housing unit" (U.S. Census Bureau, 2004a, Attachment 4, p. 3). The survey uses a probability sampling method. In addition, respondents are selected on the basis of area of residence such that the sample is representative of the state level and the nation as a whole. 57,112 housing units were interviewed from October 13-19, 2002. Respondents of the survey also answered questions concerning the demographic characteristic of each member living in their unit. The dataset contains demographic information of all 143,857 individuals residing in the selected units (U.S. Census Bureau, 2004a).

To examine the association of library service variables and the household use/non-use of the public library, this study also incorporated data from the *Public Libraries Survey* (*PLS*). The *PLS* is a part of the National Center for Education Statistics Library Statistics Program, and the data is collected through the Federal-State Cooperative System for Public Library Data. As the *CPS 2002* data was collected in 2002, the PLS data for the 2002 fiscal year was chosen for this study. This dataset includes a total of 8969 of the 9141 public libraries in the 50 states, the District of Columbia, and in the outlying areas of Guam, the Northern Mariana Islands, Palau, and the U.S. Virgin Islands (Chute et al., 2004). The dataset includes statistics on library services, such as population of a library's legal service area, and operating income and expenditures. It gives data at the national, state, and library outlet levels. This study focused on the state-level data from the 50 states and the District of Columbia.

#### 3.1. Level of analysis

The demographic, labor characteristics, and school enrollment information were collected for all 143,857 individuals residing in the selected units. The dataset was structured with one data line for each household member. The survey on library use, on the other hand, asked questions about household library use, not individual library use. For example, the question on library use was: "Has anyone in your household used a public library or bookmobile in the past year, that is since October, 2001?" (U.S. Census Bureau, 2004a, Attachment 8, p. 12). Because the dependent variable (use or non-use of the public library) is measured at the household level, the study is best analyzed using the household level as the level of analysis (Dale, Fieldhouse, & Holdsworth, 2000). As a result, this study uses household as the unit of analysis.

#### 3.2. Data preparation

To prepare the data, researchers transformed the dataset from 143,857 cases (each case representing one individual) into cases where each case represented a household. The *CPS*' definition of household focuses on the physical unit—57,112 such units were surveyed. This study, however, adopted a definition of *household* that focused on the social tie and not the physical housing unit. In this study, *household* refers to a person or group of people who reside together and consider themselves as a social unit. The rationale behind this decision was that the *CPS*' definition would include renters, who are not related to the house owner in any way other than a landlord-tenant relation, and who often have different characteristics from the house owner. Although these renters received the same household ID as the house owner, they were also interviewed. The dataset included data on the renter's characteristics, school enrollment and library use behavior. In this study, renters who were interviewed separately were given unique IDs and counted as separate households from the house owner. With these additional renters, the total number of cases for this study was 57,148.

Several variables in the dataset, such as household income, were already at a household level and required no transformation. Other variables in the dataset were recoded in SPSS so that the final dataset would include only one data line for each household. Variables can be derived in several ways, such as by summarizing the characteristics of household members and calculating the sum, maximum, or minimum value for a variable, or by choosing a reference person (such as the household head) to represent the other members of the household (Dale et al., 2000). For example, the original data set included the education level of each household member. Researchers transformed this variable into a household-level variable to show the highest level of school attainment in the household. The list of variables used in this study is provided in the next section. At the end of the recoding process, the dataset was reduced from 143,857 cases (each representing one household member) to 57,148 cases (each representing one household). The state library services data from *Public Libraries Survey* were then merged with the restructured *CPS* data. Based on the state where a household was located, the library services data of the corresponding state was added to each of the 57,148 cases.

#### 3.3. Variables

#### 3.3.1. Dependent variable

The dependent variable is a dichotomous variable at the household level, indicating the household use or non-use of the public library. In this study, households that had at least one household member who used a public library in the past year (October 2001 to October 2002) were classified as public library users. Households where no member used the public library in the last year were classified as public library nonusers.

#### 3.3.2. Independent variables

The study included 43 independent variables at the household level (Table 1). Variables were selected based on literature on public library use and on information behavior. Prominent characteristics often tested in other public use studies (such as education, income, sex, ethnicity, and marital status) were included in this research. This study also included variables that were less often examined in prior studies, such as geographic region, job employment status, and state libraries' service variables. Also included were variables that might help examine the public library use/non-use by specific groups, such as people with disabilities, ethnic minorities, recent immigrants, or the elderly. As the analysis was conducted at the household level, individual characteristics (such as cognitive or behavioral factors) were not investigated.

Household demographics, social characteristics	Household educational characteristics	Household economic, job- related characteristics	Household's use of other libraries	Household locational characteristics	Public libraries' service variables (state level)
Household size	Highest education attainment in the household	Household income	User of PreK-12 libraries	Region	Number of central libraries per 10,000 population in the state
Race/ethnicity of householder	Age 3 to 5: School attendance status	Household member owning business	User of academic libraries	Size of residing city	Number of branch libraries per 10,000 population in the state
Marital status of householder	Age 6 to 13: School attendance status	Major occupation categories of householder	User of libraries at workplace	Distance from closest public library	Libraries' total staff expenditures per state capita
Sex of householder	Age 14 to 17: School attendance status	Number of work hours per week — householder	User of other libraries (other than PreK-12, academic, workplace libraries)		Libraries' collection expenditures per state capita
Includes member with disabilities	Age 18 to 20: School attendance status	Includes household member who is looking for job			Libraries' expenditures for electronic access per state capita
Includes member who is non-U.S. citizen	Age 21 to 29: School attendance status	Household with no telephone			Number of public Internet terminals in public libraries per state capita
Includes immigrant (year of entry: before 1980)	Age 30 to 39: School attendance status	Percentage of retired persons in the household			
Includes immigrant (year of entry: 1980s)	Age 40 to 49: School attendance status	Percentage of working adults in the household			
Includes immigrant (year of entry: 1990s)	Age 50 to 59: School attendance status				
Includes recent immigrant (year of entry: 2000 or later)	Age 60 to 69: School attendance status				
	Age 70 to 79: School attendance status				
	Age 80 or above: School attendance status				

#### Table 1. Independent variables

#### 3.4. Data analysis method

Researchers used logistic regression to test the relationships between the independent variables and the household use/non-use of public libraries. Logistic regression is a variant of regression. Similar to multiple regression, logistic regression can be used in studies with continuous or categorical independent variables. Regression is used when the dependent variable is continuous, while logistic regression is suitable for a categorical dependent variable that is dichotomous in nature. Different from linear regression, logistic regression, logistic regression requires neither normally distributed data nor linear relationships between the independent and the dependent variables. Logistic regression also does not assume multivariate normality or equal variance, both of which are required for discriminant analysis. Because of these less stringent assumptions, logistic regression is increasingly applied in many research situations (Hair, Anderson, Tatham, & Black, 1998).

Before the logistic regression was conducted, the independent variables were screened for multicollinearity. Multicollinearity occurs when two or more independent variables are highly correlated and the effects of some variables might be underestimated. A general rule of thumb is that if a score on the diagnostic measure of tolerance is lower than 0.2, a multicollinearity problem could occur (Menard, 2002). In this study, the diagnostic test was conducted using the Collinearity Diagnostics function in SPSS (Leech, Barrett, & Morgan, 2005). The result showed that all 43 variables have a tolerance level higher than 0.2. Thus, all 43 variables were included in the logistic regression analysis. The Binary Logistic function offered in SPSS was used to conduct the simultaneous logistic regression analysis.

#### 4. Limitations of the study

This study is limited to civilian, non-institutional population living in households. Further study of people outside of this population (for example, those living in homes for the elderly, or homeless people) could help identify other possible under-served groups. The study's unit of analysis is also limited to the household level. Analyses at the individual level might help increase the percentage of variance explained by the logistical regression model. Individual variables such as personality, cognitive, and affective characteristics might also be a promising area to incorporate into future library use/non-use studies. In addition, future nationwide studies could compare the association of household or individual characteristics with the use/non-use of specific library services to help fine-tune and market different public library services. The scarcity of large-scale nationwide public library use data, however, hinders the expansion of research into the aforementioned topics. The Census Bureau only occasionally surveyed the public library use behavior. The *CPS Library Use* supplement used in this study, for example, was conducted in October 2002. None of the 31 *CPS* surveys released since 2002 have included the *Library Use* supplement. The availability of more recent user behavior data will be a significant factor in the future development of public library research.

# 5. Results

Of the 57,148 households, 27,511 households (48.1%) had a household member who used the public library in the past year. The following sections will present the findings on whether and how the household variables are related to the use/non-use of the public library in the past year.

#### 5.1. Summary statistics for the logistic regression model

The first research question examined whether library users and non-users differ in their characteristics. This was evaluated by examining the significance and power of the overall logistic model. All 43 variables listed in Table 1 were entered into the logistic model for analysis, and the model was significant ( $\chi^2 = 12354.76$ , df = 100, p < 0.05). The findings suggest that the independent variables, taken together, are significantly associated with the use/non-use of the public library. The amount of variances explained by the overall model is indicated by Nagelkerke  $R^2$ , a statistic comparable to the  $R^2$  in regression analysis. Its values can range from 0 to 1 (Hair et al., 1998). For this model, the Nagelkerke  $R^2$  is 260, indicating that together, the independent variables used in this analysis account for 26% of the variance in the dependent variable (i.e., household use or non-use of the public library).

#### 5.2. Contribution of each independent variable

The second research question focused on the direction and magnitude of relationships between each independent variable and the library use/non-use. In logistic regression, the *Wald* statistic is used to test whether the contribution of each independent variable is statistically significant. In this study, 34 out of the 43 variables were significant (p < 0.05). The statistically significant variables are indicated in Table 2 with \* in the next-to-last column.

Variable	В	Standar d error	Sig.		Odds ratio
Highest educational attainment in the household			0.0000	**	
Less than 1st grade [reference group]					
1st, 2nd, 3rd or 4th grade	0.15	0.50	0.7694		1.16
5th or 6th grade	0.18	0.46	0.6952		1.20
7th or 8th grade	0.25	0.44	0.5699		1.29
9th grade	0.26	0.45	0.5669		1.29
10th grade	0.62	0.44	0.1589		1.86
11th grade	0.81	0.44	0.0634		2.26
12th grade no diploma	0.94	0.44	0.0341	*	2.56
High school grad-diploma or equivalent	1.18	0.43	0.0067	**	3.24
Some college but no degree	1.65	0.43	0.0001	**	5.20
Associate degree-occupational/vocational	1.64	0.43	0.0002	**	5.18
Associate degree-academic	1.88	0.44	0.0000	**	6.53
Bachelor's degree	2.05	0.43	0.0000	**	7.74
Master's degree	2.17	0.44	0.0000	**	8.79
Professional school degree	2.31	0.43	0.0000	**	10.10

Table 2. Logistic regression results

Doctorate degree	2.28	0.44	0.0000	**	9.79
Distance from closest public library			0.0000	**	
Less than 1 mile		0.05	0.0000	**	2.30
1 or 2 miles	0.64	0.04	0.0000	**	1.90
3 to 5 miles	0.47	0.04	0.0000	**	1.60
6 to 10 miles	0.27	0.05	0.0000	**	1.31
More than 10 miles [reference group]					
Member aged 6 to 13, school attendance status			0.0000	**	
Included, attending school	0.74	0.04	0.0000	**	2.11
Included, not attending school	0.65	0.18	0.0003	**	1.91
Not included [reference group]					
User of PreK-12 libraries		0.03	0.0000	**	1.78
User of academic libraries	0.53	0.04	0.0000	**	1.69
User of other libraries	0.47	0.07	0.0000	**	1.60
Member aged 14 to 17, school attendance status			0.0000	**	
Included, attending school	0.44	0.04	0.0000	**	1.55
Included, not attending school	0.23	0.16	0.1439		1.26
Not included [reference group]					
Member aged 3 to 5, school attendance status			0.0000	**	
Included, attending school	0.36	0.05	0.0000	**	1.43
Included, not attending school	- 0.11	0.06	0.0903		0.90
Not included [reference group]					
Member aged 40 to 49, school attendance status			0.0007	**	
Included, attending school	0.36	0.10	0.0002	**	1.43
Included, not attending school	0.00	0.03	0.9964		1.00
Not included [reference group]					
Major occupation categories of householder			0.0000	**	
Managerial & professional	0.29	0.04	0.0000	**	1.33
Service occupations	0.26	0.05	0.0000	**	1.30
Production, craft, repair	0.15	0.05	0.0012	**	1.16
Farming, torestry & fishing		0.08	0.0023	**	1.26
Not working [reference group]					
Number of branch libraries per 10,000 population in the state	0.28	0.10	0.0036	**	1.32
Household income			0.0000	**	
0 to 9999 [reference group]					
10,000 to 19,999	0.11	0.05	0.0155	*	1.12
20,000 to 29,999	0.22	0.05	0.0000	**	1.25
40,000 to 39,999	0.20	0.05	0.0000	**	1.30
50,000 to 59,999	0.27	0.05	0.0000	**	1.31
60,000 to 74,999	0.28	0.06	0.0000	**	1.30
75,000 or more	0.15	0.05	0.0028	**	1.17
No answer	- 0.03	0.05	0.4610		0.97
Member aged 80 or above, school attendance status			0.0000	**	
Included, attending school	2.20	1.19	0.0649		9.03
Included, not attending school		0.05	0.0000	**	0.68
Not included [reference group]					
Member aged 30 to 39, school attendance status			0.0002	**	
Included, attending school		0.08	0.0011	**	1.29
Included, not attending school	- 0.05	0.03	0.0804		0.95
Not included [reference group]					
Includes recent immigrant (year of entry: 2000 or later)		0.08	0.0000	**	0.71
Includes member who is looking for job		0.05	0.0000	**	1.28

Race/ethnicity of householder			0.0000	**	
Non-Hispanic Black	- 0.16	0.03	0.0000	**	0.85
Non-Hispanic Native American	- 0.32	0.10	0.0011	**	0.73
Non-Hispanic Asian		0.06	0.0028	**	0.84
Hispanic		0.04	0.0005	**	0.86
Non-Hispanic White [reference group]					
Member aged 18 to 20, school attendance status			0.0000	**	
Included, attending school	0.24	0.05	0.0000	**	1.27
Included, not attending school	0.15	0.06	0.0106	*	1.16
Not included [reference group]					
User of libraries at workplace		0.05	0.0000	**	1.25
Size of residing city			0.0000	**	
Non-metropolitan	0.09	0.04	0.0458	*	1.09
100,000249,999	0.12	0.05	0.0140	*	1.13
250,000499,999	0.22	0.05	0.0000	**	1.25
500,000999,999	0.07	0.05	0.1124		1.07
1,000,0002,499,999	0.16	0.04	0.0001	**	1.18
2,500,0004,999,999	0.11	0.05	0.0209	*	1.12
5,000,000 or above [reference group]					
Marital status of householder			0.0000	**	
Married-spouse present [reference group]					
Married-spouse absent	- 0.25	0.08	0.0022	**	0.78
Widowed	- 0.23	0.04	0.0000	**	0.79
Divorced	- 0.15	0.05	0.0001	**	0.88
Separated	-0.25	0.00	0.0000	**	0.78
	0.21	0.05	0.0000		0.81
Household with no telephone	- 0.23	0.05	0.0000	**	0.80
Region			0.0000	**	
Division 1: New England [reference group]					
Division 2: Middle Atlantic	- 0.09	0.05	0.0521		0.91
Division 3: East North Central	0.12	0.05	0.0109	*	1.13
Division 4: West North Central	0.16	0.05	0.0006	**	1.17
Division 5: South Atlantic	0.04	0.06	0.5220		1.04
Division 6: East South Central	- 0.01	0.07	0.82/8		0.99
Division /: West South Central		0.06	0.1998	*	0.92
Division 0: Decific	0.12	0.05	0.0519		1.12
	0.04	0.05	0.4320		1.04
Male householder	- 0.14	0.02	0.0000	**	0.87
Member aged 21 to 29, school attendance status			0.0000	**	
Included, attending school	0.08	0.05	0.1189		1.09
Included, not attending school	- 0.13	0.03	0.0001	**	0.88
Not included [reference group]					
Member aged 70 to 79, school attendance status			0.0060	**	
Included, attending school	- 0.14	0.54	0.7905		0.87
Included, not attending school	- 0.13	0.04	0.0014	**	0.88
Not included [reference group]					
Includes member with disabilities		0.03	0.0004	**	0.88
Household member own business		0.03	0.0035	**	1.09
Household size		0.02	0.0000	**	1.07
Libraries' collection expenditures per state capita		0.02	0.0041	**	1.05
Number of public Internet terminals in public libraries per state capita		0.02	0.0301	*	1.03
Percentage of working adults in the household		0.00	0.0000	**	1.00

Number of work hours per week—householder	0.00	0.00	0.0017	**	1.00
Percentage of retired persons in the household		0.00	0.0317	*	1.00
Includes immigrant (year of entry: 1990s)	- 0.10	0.06	0.0792		0.90
Includes immigrant (year of entry: before 1980)	0.02	0.05	0.6668		1.02
Includes immigrant (year of entry: 1980s)		0.06	0.1513		1.08
Includes member who is non-U.S. citizen	- 0.11	0.06	0.0633		0.90
Number of central libraries per 10,000 population in the state	- 0.02	0.04	0.5394		0.98
Libraries' total staff expenditures per state capita		0.00	0.0854		1.01
Libraries' expenditures for electronic access per state capita	0.01	0.01	0.4483		1.01
Member aged 50 to 59, school attendance status			0.0563		
Included, attending school	0.29	0.15	0.0515		1.33
Included, not attending school	- 0.03	0.03	0.2394		0.97
Not included [reference group]					
Member aged 60 to 69, school attendance status			0.1277		
Included, attending school	0.24	0.33	0.4561		1.28
Included, not attending school	- 0.06	0.03	0.0638		0.94
Not included [reference group]					
Constant	0.17	0.49	0.7348		1.18

The odds ratio is used to examine how changes in each independent variable (and the levels within the variable) affect the direction and magnitude of changes in the dependent variable while other factors remain constant (Table 2). The interpretation of the odds ratio (OR) is as follows: *ceteris paribus*, a household with a particular characteristic is more likely to be a library user if that characteristic has an odds ratio higher than 1 (OR > 1). The higher the odds ratio, the higher the chance they will belong to the library user group. Conversely, a household is less likely to be a public library user if that characteristic has an odds ratio smaller than 1 (OR < 1). For example, in the "highest level of education" variable, the "doctorate degree" category has an odds ratio of 9.79. It means that a household with this characteristic is more likely to be a library user. Specifically, the chance that this household will use a public library is 9.79 times higher than the household in the reference category (i.e., household whose members have less than a 1st grade education level). The odds ratios of the independent variables are listed in Table 2.

# 6. Discussion

The study found a statistically significant association between the whole group of variables and the household use/non-use of the public library. When the contribution of each independent variable was analyzed, 34 of the 43 variables were found to be significant. This finding indicates that the public library users and the non-users in this study differed in many characteristics (Table 3).

# Table 3. Summary of findings

# Households are more likely to be library users if they have one or more of the following characteristics, and vice versa:

# Household demographics and social characteristics

- Larger household size
- Householder is not of ethnic minority
- Householder is married, and spouse is present
- Female householder
- Household that do not include a member with disabilities
- Household members are not recent immigrants

# Household education characteristics

- Household with higher education attainment
- Household with members aged 3 to 20, or 30 to 49, who are attending schools
- Household with members aged 6 to 13, or 18 to 20, who are not attending schools
- Household without members aged 21 to 29, or 70 or above, who are not attending schools

# Household economic and job-related characteristics

- Household with high but not the highest income
- Household with working householder, especially in managerial or professional occupation
- Household members own business
- Household members looking for job
- Householder did not work too many hours per week
- Household with a telephone
- Higher percentage of retired person
- Lower percentage of working adults

# Household use of libraries

• Household using PreK-12 libraries, academic libraries, libraries at work, or other libraries

# Household locational characteristics

- Household that is close by a public library
- Household that resides in a medium size city
- Household in the East North Central, West North Central, or Mountain states

# Public Libraries services variables (state level)

- Household in states where there are more branch libraries per 10,000 population
- Household in states with higher library collection expenditures per state capita
- Household in states with higher number public Internet terminals per state capita

This study's multivariate analysis contributes to addressing the problem caused by highly correlated variables.

It helps identify and mitigate possible spurious effects. The results show that education attainment

contributes the most to the model, which concurs with previous studies.

While there is general agreement concerning the positive relation of education level and public library use, little agreement exists regarding the role of income level. This uncertainty is partly related to the high correlation between income and education levels (e.g., Kronus, 1973; Zweizig & Dervin, 1977). This study's findings indicate that income is indeed highly correlated with education (Pearson r = 0.362, p < 0.05). The logistic regression results show that after controlling all other variables in the study, income does not contribute as strongly as education level, yet income is still significantly associated with library use/non-use.

Based on the odds ratios of the household income variable, the findings suggest that the relationship between household income and household use/non-use of the public library is mildly curvilinear (Fig. 1). This curvilinear relationship might have added to the difficulty in evaluating its effect. Multiple regression or discriminate analysis is often used in multivariate studies of the public library, but these two methods assume that the independent and dependent variables have a linear relationship. Logistic regression, on the other hand, does not use this assumption. The curvilinear relationships found in the income in this study highlight the need for future studies to test and take into account possible non-linear relationships. This study suggests that after controlling for other variables, households with middle income level are more likely to be library users than households with very low or very high income. The result concerning the less-advantaged, low-income population is especially worth noting. It stresses the importance of continuous efforts to ensure that the needs of this population are well-served.



Fig. 1. Odds ratio for household income. Note: Reference group: Household with income from \$0 to \$9999.

The study found a significant association between distance from closest public library and library use/non-use. Several studies have found distance or perceived distance to be significant in library use/non-use (D'Elia, 1980; Lange, 1988; University of Michigan Survey Research Center, 1948). Palmer (1981) offered a review of the literature on this topic. Nevertheless, other studies found distance to be a non-significant variable (e.g., Evans, 1970; Marchant, 1994). A particularly interesting finding of this study is the high strength of this variable (for households living less than 1 mile from a library, OR = 2.30). A possible explanation for this finding is that the changing information environment might have made proximity a more important criterion. Studies on information behavior have revealed that convenience and accessibility are key criteria in selecting information sources (O'Reilly, 1982). Especially in this decade, an increased expansion of information channels and sources are available to users. Many users enjoy the accessibility of the Internet and look there for information. Since there are more accessible information sources available outside of the public library, the convenience/accessibility criterion (e.g., distance from the library) might play a bigger role in the use/non-use behavior than before. This hypothesis is worth further exploration. Findings on this issue can have practical implications on the future planning of the public library, such as whether to build more branch libraries or to provide more services online.

In previous studies, the relation between school attendance and public library use was seldom tested directly. Instead, school attendance was often discussed in conjunction with age, especially in the discussion of library use by children and young adults. Berelson (1949), for example, suggested that "the predominance of children and young people in the clientele of the public library suggests another important fact, namely, that the use of the library falls off sharply at the school-leaving age" (p. 21). It is uncertain whether public library uses are associated with age, school attendance, or both. To address this issue, this study created age group-school attendance status variables for eleven age groups (e.g., age 3 to 5, school attendance status). These variables noted whether a household had a member in a particular age group and whether that member is attending school. The results of the 11 variables can be found in Table 2 and are presented together in Fig. 2 for comparison. The findings show that households with members who are not attending schools and are aged 6 to 13 or 18 to 20 are more likely to be library users than household without members in the respective age groups. This finding is similar to previous studies (Berelson, 1949; Collins & Chandler, 1997; Marchant, 1994; Westin & Finger, 1991). In contrast with previous studies, this study distinguished age from school attendance. It found that the presence of members in those age groups itself can increase the odds of library use. The study also found that household members age 21 to 29 are less likely to use public libraries, perhaps because the transition from school life to work or family life may consume more of people's time and effort, contributing to a decrease in library use. All other variables held constant, households with members in their 70s or 80s are also less likely to be library users than those households without such members. Public library

services for senior citizens needs further study, especially when it is projected that by 2050, about one-fifth of the U.S. population will be over 65 years of age (U.S. Census Bureau, 2004b).



Fig. 2. Odds ratios for the age group-school attendance status variables. Notes: Significance level at 0.05. Odds ratio higher than 1 indicates higher likelihood of public library use when compared to the reference category (i.e., households who did not have member in the respective age group).

If a household member is attending school, the odds of library use increase. For all but one age group, households with a member attending school have a higher odds ratio than households with a member in that age group, not attending school. For example, households with members aged 3 to 5, 30 to 39, or 40 to 49, attending school, are all more likely to be library users. Those with members in the respective age groups who are not attending school are not statistically significant. The pattern for those that attend schools is more akin to a bimodal pattern than a curvilinear pattern. As Zwezig and Dervin (1977) reported in their literature review, the results on age have been mixed. This study suggests that differentiating age from school attendance might help clarify the role of age in library use. The results for both the attending-school and not-attending-school groups exhibit a non-linear pattern, thus further highlighting the need to check for non-linear relationships in future research.

The higher odds of library use by households with members attending school lead to another question that is less investigated—the relationship between the use of other types of libraries and use of the public library. School attendance means access to other types of libraries, such as PreK-12 or academic libraries. To a

certain extent, these libraries could be considered a substitute for the public library and might lower the likelihood of public library use. However, this study suggests that the use of other types of libraries might not diminish the use of public libraries. On the contrary, households that use other libraries are also more likely to be public library users. The associations between the uses of PreK-12 and academic libraries with the uses of public libraries are practically strong (OR = 1.78 and 1.69, respectively). While future studies are needed to identify the purpose of using these different types of libraries, the findings suggest that the libraries complement rather than compete with each other. Collaboration between public libraries and PreK-12/academic libraries (Fitzgibbons, 2000) and further efforts to promote public libraries in schools should be encouraged.

Library service variables that have been less examined in previous studies of public library use/non-use were found to be statistically significant in this study. The results suggest that library use/non-use research could benefit from including supply-side variables. For example, households in states with higher collection expenditure per state capita or with more public Internet terminals per state capita are more likely to be library users (OR = 1.05 and 1.03, respectively). Households in states where there are higher numbers of branch libraries per 10,000 population are also more likely to be library users (OR = 1.32). The number of central libraries per 10,000 population, on the other hand, was not found to be statistically significant. These findings on branch libraries further suggest the importance of accessibility to public library use/non-use. The current analysis was conducted using the library service variables at the state level, as the household data in *CPS* indicated the state where the households reside but did not include full information on the county/district. Future research of service variables at a county or neighborhood level could provide more detailed information.

Ethnicity is statistically significant (p < 0.05). The findings differ from previous studies that found race/ethnicity to be non-significant (Evans, 1970; Kronus, 1973; Zweizig, 1973; Zweizig & Dervin, 1977). One of the possible differences between this study and previous studies is the difference in sample size. Even for nationwide studies, the sample size is typically below 2000. The number of respondents belonging to each ethnic minority group might be relatively small. The national survey published by ALA, for example, does not include any Native Americans or Asian Americans (Westin & Finger, 1991). For the current study, 44,582 of the householders are Non-Hispanic White, 5755 are Non-Hispanic African American, 4351 are Hispanic, 1875 are Non-Hispanic Asian American, and 585 are Non-Hispanic Native American. This study finds that all four ethnic minority groups have an odds ratio less than one. It shows that even after holding other variables constant, households whose head is of an ethnic minority are still less likely to be library users when compared to the families of Non-Hispanic White householders. The finding is worth noting, as the U.S. population is increasingly diverse. By 2050, ethnic minorities are projected to reach 50% of the total U.S. population (U.S. Department of Commerce, Minority Business Development, 1999). In order to serve the increasingly diverse communities, a better understanding of the interaction between race/ethnicity and the use/non-use of public libraries is needed. Research on the information behavior of, and types of services used by, different ethnic groups, such as those recommended by Koontz, Jue, and Lance (2005) would contribute to better services. More efforts and discussion concerning multiculturalism and diversity in the LIS profession would also be beneficial.

# 7. Conclusion

The study findings indicate that a number of less-advantaged groups are still more likely to be non-users of public libraries. While people might choose not to use public libraries because alternative resources are available, it should be noted that many of these non-users' groups have less access to information resources in general (Duke, 2000; Evans, 2004; Roscigno et al., 2006; Spink & Cole, 2001]). To narrow such information divides, public libraries should proactively reach out to these disadvantaged groups. Building more branch libraries in areas closer to these groups could be beneficial, as the distance to the library seems to have a strong impact on the use of the public library. In addition, public libraries should strive to provide collections and services that are relevant to these groups. They could also raise awareness of different library services available to the disadvantaged groups through targeted marketing. Many LIS professionals have highlighted the need to better serve the less-advantaged population. However, such endeavors are often impeded by limited library budgets. The persistent funding problems experienced by library systems throughout the nation are alarming (Davis, 2006). Narrowing the information divides requires serious effort and commitment from the LIS professionals, but libraries also need resources to be made available to aid such endeavors. If the nation is determined to overcome such information inequity in its population, steadfast commitment and support from all levels of governments and different sectors of the society are of paramount importance.

# Acknowledgments

The authors wish to thank Dr. Louise Robbins and Dr. Ethelene Whitmire for their helpful suggestions on an earlier version of the paper.

# References

- Agada, J. (1999). Inner-city gatekeepers: An exploratory survey of their information use environment. *Journal* of the American Society for Information Science, 50, 74–85.
- Agosto, D. E., & Hughes-Hassell, S. (2005). People, places, and questions: An investigation of the everyday life information-seeking behaviors of urban young adults. *Library & Information Science Research*, 27, 141–163.

Berelson, B. (1949). The library's public: A report of the public library inquiry. New York: Columbia University Press.

- Bertot, J. C., McClure, C. R., Jaeger, P. T., & Ryan, J. (2006). Public libraries and the Internet 2006: Study results and findings. Retrieved January 10, 2007, from http://www.ii.fsu.edu/projectFiles/plinternet/2006/2006 plinternet.pdf
- Bill & Melinda Gates Foundation. (2004). Toward equality of access: The role of public libraries in addressing the digital divide. Retrieved January 10, 2007, from <a href="http://www.gatesfoundation.org/nr/Downloads/libraries/uslibraries/reports/TowardEqualityofAccess.pdf">http://www.gatesfoundation.org/nr/Downloads/libraries/uslibraries/reports/TowardEqualityofAccess.pdf</a>
- Bishop, A., Tidline, T., Shoemaker, S., & Salela, P. (1999). Public libraries and networked information services in low-income communities. *Library & Information Science Research*, 21, 361-390.
- Campbell, A., & Metzner, C. A. (1952). *Public use of the library and of other sources of information*. Ann Arbor: University of Michigan (rev. ed.).
- Chatman, E. A. (1985). Information, mass media use, and the working poor. Library & Information Science Research, 7, 97-113.
- Chen, C. C., & Hernon, P. (1982). Information seeking: Assessing and anticipating user needs. New York: Neal-Schuman.
- Chute, A., Kroe, E., O'Shea, P., Craig, T., Freeman, M., Hardesty, L., et al. (2004). Public libraries in the United States: Fiscal year 2002. Retrieved January 3, 2007, from National Center for Education Statistics Web site: <u>http://nces.ed.gov/pubs2005/2005356.pdf</u>
- Chute, A., Kroe, E., O'Shea, P., Craig, T., Freeman, M., Hardesty, L., et al. (2006). Public libraries in the United States: Fiscal year 2004. Retrieved January 3, 2007, from National Center for Education Statistics Web site: <u>http://nces.ed.gov/pubs2006/2006349.pdf</u>
- Collins, M. A., & Chandler, K. (1997). Use of public library services by households in the United States: 1996. Retrieved January 3, 2007, from National Center for Education Statistics Web site: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=97446
- Cooper, S. M., Bolt, N., Lance, K. C., & Webster, L. (1993). *Community analysis methods and evaluative options: The CAMEO handbook*. Retrieved January 3, 2007, from <u>http://skyways.lib.ks.us/pathway/cameo/</u>
- Dale, A., Fieldhouse, E., & Holdsworth, C. (2000). Analyzing census microdata. London: Arnold.
- Davis, D. M. (2006). Funding issues in U.S. public libraries, fiscal years 2003–2006. Retrieved January 11, 2007, from http://www.ala.org/ala/ors/reports/FundingIssuesinUSPLs.pdf
- D'Elia, G. (1980). The development and testing of a conceptual model of public library use behavior. *Library Quarterly, 50*, 410–430.
- D'Elia, G. (1981). Procedure for developing a typology of adult users of the public library. *Library Research*, *3*, 123–140.
- D'Elia, G. (1993). The roles of the public library in society: The results of a national survey, final report. Evanston: Urban Libraries Council.
- D'Elia, G., Jörgensen, C., Woelfel, J., & Rodger, E. J. (2002). The impact of the Internet on public library use: An analysis of the current consumer market for library and Internet services. *Journal of the American Society for Information Science & Technology*, 53, 802–820.
- Dervin, B., & Nilan, M. (1986). Information needs and uses. *Annual Review of Information Science and Technology*, 21, 3–33.

- Duke, N. K. (2000). For the rich it's richer: Print experiences and environments offered to children in very low- and very high-socioeconomic status first-grade classrooms. *American Educational Research Journal*, 37, 441–478.
- Evans, C. (1970). Middle class attitudes and public library use. Littleton: Libraries Unlimited.
- Evans, G. W. (2004). The environment of childhood poverty. American Psychologist, 59(2), 77-92.
- Fitzgibbons, S. A. (2000). School and public library relationships: Essential ingredients in implementing educational reforms and improving student learning [Electronic version]. School Library Media Research, 3. Retrieved January 15, 2007, from <u>http://www.ala.org/ala/aasl/aaslpubsandjournals/slmrb/slmrcontents/volume32000/relationships. htm</u>
- Fox, S. (2005). *Digital divisions*. Retrieved September 28, 2007, from http://www.pewinternet.org/pdfs/PIP Digital Divisions Oct 5 2005.pdf
- Fox, S., & Livingston, G. (2007). Latinos online. Retrieved September 28, 2007, from http://www.pewinternet.org/pdfs/Latinos Online March 14 2007.pdf
- Gallup Organization. (1976). The role of libraries in America. Frankfort: Kentucky Department of Libraries and Archives.
- Gallup Organization. (1978). Book reading and library usage: A study of habits and perceptions. Princeton: Gallup Organization.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River: Prentice Hall.
- Koontz, C. M., Jue, D. K., & Lance, K. C. (2005). Neighborhood-based in-library use performance measures for public libraries: A nationwide study of majority–minority and majority white/low income markets using personal digital data collectors. *Library & Information Science Research*, 27, 28–50.
- Kronus, C. I. (1973). Patterns of adult library use: A regression and path analysis. *Adult Education*, 23, 115–131.
- Lange, J. M. (1988). Public library users, nonusers, and type of library use. *Public Library Quarterly*, 8(1/2), 49–67.
- Leech, N. L., Barrett, K. C., & Morgan, G. A. (2005). SPSS for intermediate statistics: use and interpretation (2nd ed.). Mahwah: Lawrence Erlbaum.
- Lynch, M. J. (1997). Using public libraries: What makes a difference? American Libraries, 28, 64-66.
- Madden, M. (1979). Library user/nonuser lifestyles. American Libraries, 10, 78-81.
- Marchant, M. P. (1994). Why adults use the public library: a research perspective. Englewood: Libraries Unlimited.
- McCook, K. d. l. P., & Geist, P. (1994). Toward a just and productive society: An analysis of the recommendations of the White House Conference on Library and Information Services. Washington, D.C.: National Commission on Library and Information Science.
- Menard, S. W. (2002). Applied logistic regression analysis (2nd ed.). Thousand Oaks: Sage.
- Mishel, L., Bernstein, J., & Allegretto, S. (2006). The state of working America 2006/2007. Ithaca: ILR Press.
- Online Computer Library Center. (2004). 2004 information format trends: Content, not containers. Retrieved September 15, 2007, from OCLC Web site: <u>http://www.oclc.org/reports/2004format.htm</u>
- O'Reilly, C. A. (1982). Variations in decision makers use of information-sources—The impact of quality and accessibility of information. *Academy of Management Journal*, 25, 756–771.

- Palmer, E. S. (1981). The effect of distance on public library use: A literature survey. *Library Research*, *3*, 315-354.
- Powell, R. R. (1984). Library use and personality: The relationship between locus of control and frequency of use. Library & Information Science Research, 6, 179–190.
- Powell, R. R., Taylor, M. T., & McMillen, D. L. (1984). Childhood socialization: Its effect on adult library use and adult reading. *Library Quarterly*, 54, 245–264.
- Rees, M. B., & Paisley, W. J. (1968). Social and psychological predictors of adult information seeking and media use. *Adult Education Journal*, 19, 11–29.
- Roscigno, V. J., Tomaskovic-Devey, D., & Crowley, M. (2006). Education and the inequalities of place. Social Forces, 84, 2121-2145.
- Scheppke, J. (1994). Who is using the public library? Library Journal, 119(17), 34-37.
- Spink, A., & Cole, C. (2001). Information and poverty: Information-seeking channels used by African American low-income households. *Library & Information Science Research*, 23, 45-65.
- U.S. Census Bureau. (2004a). Current population survey, October 2002. School enrollment /Library use supplement file. Technical documentation. Retrieved March 2, 2007, from Census Bureau Web site: http://www.census.gov/apsd/techdoc/cps/cpsoct02.pdf
- U.S. Census Bureau. (2004b). Projected population of the United States, by race and Hispanic Origin: 2000 to 2050. Retrieved March 2, 2007, from Census Bureau Web site: <u>http://www.census.gov/ipc/www/usinterimproj/</u>
- U.S. Department of Commerce Minority Business Development Agency. (1999). Dynamic diversity: Projected changes in U.S. race & ethic composition 1995 to 2050. Retrieved March 2, 2007, from Minority Business Development Agency Web site: <a href="http://www.mbda.gov/documents/unpubtext.pdf">http://www.mbda.gov/documents/unpubtext.pdf</a>
- University of Michigan Survey Research Center. (1948). The public library and the people: A national survey done for the Public Library Inquiry. Ann Arbor: University of Michigan.
- Van House, N. A. (1983). A time allocation theory of public library use. Library & Information Science Research, 5, 365-384.
- Westin, A. F., & Finger, A. L. (1991). Using the public library in the computer age: Present patterns, future possibilities. Chicago: American Library Association.
- Wilson, T. D. (1999). Models in information behavior research. Journal of Documentation, 55, 249-270.
- Zhu, X. D. (2007). Growing wealth, inequality, and housing in the United States. Retrieved September 15, 2007, from the Joint Center for Housing Studies, Harvard University Web site: http://www.jchs.harvard.edu/publications/markets/w07-1.pdf
- Zweizig, D. (1973). Predicting amount of library use: An empirical study of the role of the public library in the life of the adult public. Doctoral dissertation, Syracuse University.

Zweizig, D., & Dervin, B. (1977). Public library use, users, uses: Advances in knowledge of characteristics and needs of the adult clientele of American public libraries. *Advances in Librarianship*, 7, 232–255.