The Impact of Library Resource Utilization on Undergraduate Students' Academic Performance: A Propensity Score Matching Design

Felly Chiteng Kot and Jennifer L. Jones

This study uses three cohorts of first-time, full-time undergraduate students (N=8,652) at a large, metropolitan, public research university to examine the impact of student use of three library resources (workstations, study rooms, and research clinics) on academic performance. To deal with self-selection bias and estimate this impact more accurately, we used propensity score matching. Using this unique approach allowed us to construct treatment and control groups with similar background characteristics. We found that using a given library resource was associated with a small, but also meaningful, gain in first-term grade point average, net of other factors.

mid budget cuts and legislative pressure, academic institutions are increasingly seeking ways to foster student success. Colleges and universities often create and/or expand support services and programmatic interventions, with the hope that these services will yield positive returns, such as higher student persistence rates and better grades. Student support services, in turn, are increasingly asked to demonstrate their worth in terms of contributing to the achievement of institutional outcomes. Academic libraries are not an exception.[1] As Oakleaf noted, "[l]ibrarians are increasingly called upon to document and articulate the value of academic and research libraries and their contribution to institutional mission and goals."[2]

Georgia State University Library has made a conscious effort over the past nine years to measure resource and service usage, along with user satisfaction and awareness levels, in an overall effort toward continuous improvement. This effort, which has included regular surveys, focus groups, usability studies, and other traditional assessment methods, has been helpful to the library but has "not [measured] the impact of the library on [users'] success."[3] Recently the library, in collaboration with Georgia State's Office of Institutional Effectiveness, has focused on assessing the impact of the library on student academic achievement. This effort has stemmed from the 2012 com-

Felly Chiteng Kot is Institutional Research Analyst for the Office of the Provost, Nazarbayev University, Kazakhstan, e-mail: felly.chiteng@nu.edu.kz; Jennifer L. Jones is Assessment & User Experience Librarian in the University Library at Georgia State University; e-mail: jlink@gsu.edu. © 2015 Felly Chiteng Kot and Jennifer L. Jones, Attribution-NonCommercial (http://creativecommons.org/licenses/by-nc/3.0/) CC BY-NC

mitment by Georgia Governor Nathan Deal to the Complete College Georgia program. Governor Deal pledged that Georgia's postsecondary institutions would confer 250,000 additional certifications and degrees beyond the current numbers by the year 2020.[4] In response, Georgia State University placed even greater emphasis on improving student retention and increasing the numbers of associates' and bachelors' degrees conferred each year, among other factors. The university has focused on implementing best practices intended to improve academic outcomes; as such, there is an interest in assessing how campus resources impact student success.

Georgia State University Library averages 7,000 visits by 4,500 unique visitors each day during an average semester. Each year the library records hundreds of thousands of library workstation logins, tens of thousands of group study room reservations, and hundreds of attendees at instruction sessions and workshops. The library is a busy place, and we wondered what impact, if any, student use of library services and resources had on their academic performance. First-year grades are considered to be the "single best predictors of student persistence."[5] Therefore, we determined that investigating the impact of library resource utilization on first-term grade point average (GPA) would be the place to start. In this study, we investigated the question: How does using library resources and services impact undergraduate students' academic performance in their first term?

Literature Review

The literature on assessment in academic libraries is broad and impressive, spanning decades. However, until recently, assessment has relied mostly on surveys, focus group interviews, usability testing, space studies, door counts, questionnaire responses, and the like. As Wong and Webb noted, "none of these assessment methods can measure the impact of libraries on student learning outcomes."[6] A shift in assessment focus seems to have occurred since the 2010 publication of *The Value of Academic Libraries: A Comprehensive Research Review and Report* (*VAL Report*) by the Association of College and Research Libraries. The report served as an assessment wake-up call for libraries to move from traditional reports of outputs to reports of the measurable impact they have on their respective campuses, particularly in areas such as student retention and engagement and faculty teaching and research.[7] Author Megan Oakleaf suggests a number of ways libraries can begin providing evidence of impact, including "[investigating] correlations between student library interactions and their GPA" and "demonstrating the library's role in retaining students until graduation."[8]

The literature review conducted for the *VAL Report* is exhaustive and useful in informing opportunities for libraries pursuing impact studies. Therefore, rather than focus on these same studies, we instead focus the present review on the small, but growing, body of research published simultaneous to or since the *VAL Report*. Some of these studies have been conducted at academic institutions overseas; others, at academic institutions in the United States.

Among the studies conducted overseas, Wong and Webb used a sample of more than 8,000 students who had graduated from Hong Kong Baptist University to examine the correlation between the number of books and audiovisual materials checked out during the course of the student's study program and the student's graduation GPA.[9] This study found that use of books and audiovisual materials was positively correlated with graduation GPA in 65 percent of the 48 subgroups (based on student major and level of study) examined. In a follow-up study, Wong and Cmor used the same sample and examined whether participation in library instruction workshops was positively correlated with graduation GPA.[10] They found that programs that offered more library sessions to students also tended to show a positive correlation between

student attendance at library sessions and graduation GPA. For example, only 15 percent of the programs that offered one library workshop and only 22 percent of the programs that offered two workshops to their students showed a positive correlation between workshop attendance and graduation GPA. In contrast, around 50 percent of the programs that offered three or four workshops showed a positive correlation between workstation attendance and graduation GPA.

The University of Wollongong Library (Australia) developed a database and reporting system as "a cost effective and sustainable way of collecting information on [the library's] impact on client outcomes."[11] The end product, the Library Cube, merges library use data and student demographic and academic performance data and allows for an assessment of the relationship between library usage and student performance.[12] A sample finding from the Library Cube revealed "a very strong nonlinear correlation between average usage of resources and average student marks."[13] Students who borrowed books and used electronic resources were found to have higher grades than those who did not.

A study conducted at Huddersfield University (U.K.) used a sample of more than 20,000 first- through fourth-year students and investigated library visits, use of electronic resources, and book loans.[14] This study found some indication that students who used more electronic resources and those who borrowed more books tended to have better grades. The University of Huddersfield study was later expanded to other U.K. universities. This follow-up study used a sample of more than 33,000 undergraduate students from eight U.K. universities and found a positive relationship between library resource utilization (access to electronic resources and book loans) and degree attainment.[15] This relationship held true collectively across institutions and, for institutions providing loan and electronic resource data, at the institution level.[16]

Some of the most recent studies have been conducted at U.S. academic institutions. For instance, researchers at Samford University examined the correlation between access to e-books, e-journals, online databases, and electronic reference works and the GPA of freshmen, sophomores, juniors, and seniors.[17] This study found a positive, weak to moderate, and statistically significant correlation between the two variables and across the four class levels. Researchers from the University of Minnesota sought to expand the scope of previous studies by examining multiple library resources and services, rather than just focusing on one or two resources.[18] This study used a sample of more than 5,300 first-year, nontransfer undergraduate students and included 13 library access variables, including material loans and renewals; on- and off-campus electronic resource logins; workstation logins; and library workshop registrations. After controlling for student demographic characteristics, academic background, and campus experiences, the study found that students who used the library at least once, regardless of the resource or service, during their first semester had a higher first-term GPA compared to students who did not.[19] Additionally, students who used the library at least once had a higher fall-to-spring semester retention compared to their peers who did not.[20] This study also identified a differential association between the type of library service or resource used and the outcomes of interest. For example, four resources (workstation use, online database access, electronic journal access, and book loans) were found to be related to term GPA and two resources (workshop attendance and online database access) to retention.

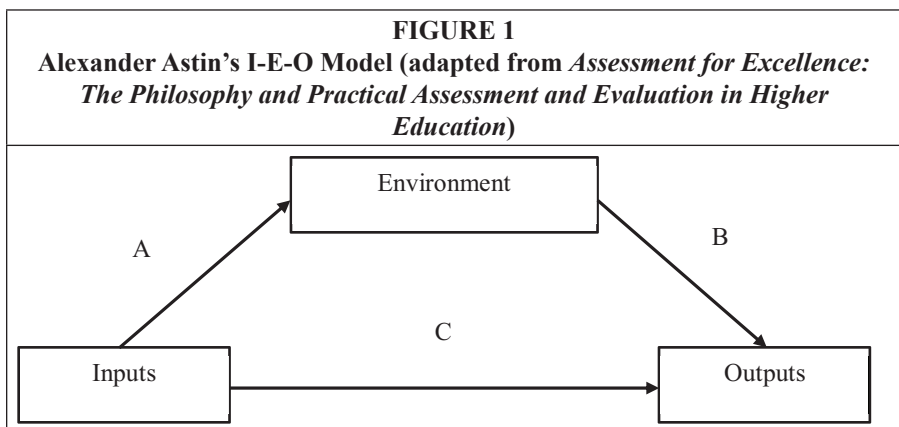With the exception of the University of Minnesota study, which used regression analysis and controlled for students' background characteristics, the other studies reviewed merely focused on bivariate correlation between use of library resources and students' academic outcomes. A consistent shortcoming in previous studies that examine the relationship between library use and student achievement is that these

studies did not take into consideration the fact that a variety of factors may contribute to students' decisions to use library resources and that these factors may be, in turn, related to the student outcomes of interest. If this were the case, then the estimated "impact" of library resources on student outcomes could, in fact, reflect (at least partially) the relationship between student characteristics and academic outcomes. In other words, the estimated relationship or impact could be biased. In the present study, we addressed this issue by using propensity score matching and constructing treatment and control groups in a way that attempts to mimic a randomized experiment. We used this approach to examine the impact of student use of three library resources (workstations, study rooms, and research clinics) on first-term GPA. In addition to adjusting for a variety of student characteristics, we computed the average treatment effect (ATE) corresponding to each library resource examined. For each ATE, we also computed the corresponding effect size to assess the practical, rather than simply the statistical, significance of the estimated impact.

Conceptual Framework

In this study, we used Astin's input-environment-output (I-E-O) model, represented by figure 1, as our conceptual framework.[21] Astin conceptualized the college as comprising three components: student inputs, the college environment itself, and student outputs. According to Astin, inputs are personal qualities that the student brings to college, whereas the environment consists of the student's actual experiences during his/her college education. The outputs consist of the student's developmental aspects that the college seeks to influence. In the present study, inputs included students' demographic characteristics and academic preparation. The primary environmental experience of interest was students' use of library resources. Secondary environmental experiences included the student's college or school, credit load, living arrangement, financial situation, and participation in Freshman Learning Communities (FLCs). The output of interest was students' academic performance in the first term.

According to Astin, analysis of the effect of environmental experiences on outputs (arrow B) is the main concern of the research on the impact of college. This is because environments can be modified to offer students a better experience and enhance their academic performance or progress. According to Astin, outputs are affected by both inputs and environmental experiences. As Astin's model further shows, inputs are almost always related to environmental experiences. This therefore presents an analytical challenge: "any observed relationship between environments and outcomes

---

FIGURE 1
Alexander Astin's I-E-O Model (adapted from *Assessment for Excellence: The Philosophy and Practical Assessment and Evaluation in Higher Education*)

might well reflect the effects of inputs rather than the actual effects of environments on outcomes."[22]

In fact, previous studies that examined the relationship between library resources and academic outcomes failed to take into account that students' use of library resources may be a function of a number of factors and that these factors, in turn, may be related to student academic outcomes. In other words, self-selection bias is a threat in studies that examined the relationship between student use of library resources and academic outcomes. This is because, in general, students make a decision to use or not use library resources. Consequently, students who use library resources may differ systematically from those who do not use these resources. To deal with self-selection, we used a quasi-experimental design to eliminate, or at least substantially decrease, the relationship between inputs (student characteristics) and the primary environmental variable of interest (use of library resources). This approach allowed us to estimate the impact of library resources on academic outcomes more accurately.

Methods
Sample and Data Sources
The sample for this study comprised 8,652 first-time, full-time freshmen who matriculated in fall 2010, fall 2011, and fall 2012 at Georgia State University—a large, metropolitan, public research university with a diverse student body. Of the student sample used, 55 percent were female, 35 percent Black, 33 percent White, 17 percent Asian, and 9 percent Hispanic. This study used two data sources. Georgia State University Library provided data on students' utilization of library workstations, group study room reservations, and research clinic attendance. The library requires that students log in to library computer workstations using their campus usernames, and login data are collected through LabStats computer lab management software. The library also provides access to 60 group study rooms that must be reserved through an online reservation system, which also requires campus username authentication. Students who attended research clinics, one-hour classes on various research-related topics, recorded their attendance by swiping their campus ID cards through magnetic swipe readers. The Office of Institutional Effectiveness merged library use data with student background characteristics and academic records extracted from the university data warehouse.

Variables
The output (or independent variable) of interest was students' academic performance, measured by first-term GPA. The environmental variable of interest was use of library resources, which included workstations, study rooms, and research clinics. Workstation usage was measured by the number of times the student logged in to library workstations during his/her first fall term on campus. To take into account the significant variations in the number of times students used library workstations, we created four separate dummy variable indicators and used each in a separate analysis. These workstation-usage indicators were as follows:
1. Whether or not the student logged in to workstations at least once.
2. Whether or not the student logged in to workstations at least five times.
3. Whether or not the student logged in to workstations at least 10 times.
4. Whether or not the student logged in to workstations at least 20 times.

Study room usage was a binary variable that indicated whether or not the student reserved a study room at least once in his/her first term. Finally, research clinic attendance measured whether or not the student attended a research clinic at least once in his/her first term. Each of the six "treatment" indicators was the focus of a separate analysis.

This study included various control variables (both input and environmental). Input variables included students' demographic characteristics and academic preparation. Demographic characteristics included the student's sex, race/ethnicity, citizenship, age at matriculation, and the matriculation term. Academic preparation included the student's high school GPA, SAT math score, and SAT verbal score, as well as an indicator of whether the student transferred any Advanced Placement (AP) credits. Other environmental variables used as control variables included the student's college or school, the number of credits taken in the first term, whether the student lived on campus, whether the student participated in a Freshman Learning Community (FLC), and the student's level of unmet financial need. Table 1 provides the frequency distribution of the student sample on categorical variables, and table 2 gives descriptive statistics on continuous and discrete variables.

Research Design and Data Analysis

Students who used library resources (as measured by the six indicators defined earlier) may have differed systematically from those who did not use these resources. Therefore, any estimate of the impact of library resources may be biased, particularly if there are variables that predict both student use of library resources and student academic performance. For example, it may be possible that female students use a

TABLE 1			
Frequency Distribution of the Sample (Categorical Variables)			
Variable	Levels	Frequency	Percent
Workstation Usage	Did not use workstations	2,579	29.81
	Used workstations at least once	6,073	70.19
	Used workstations at least 5 times	3,405	39.36
	Used workstations at least 10 times	2,136	24.69
	Used workstations at least 20 times	982	11.35
Study Room Usage	Did not use study rooms	5,186	89.09
	Used study rooms at least once	635	10.91
Research Clinics Attendance	Did not attend research clinics	2,877	92.15
	Attended research clinics	245	7.85
Gender	Female	4,765	55.07
	Male	3,887	44.93
Race/Ethnicity	Asian	1,486	17.18
	Black	3,042	35.16
	Hispanic	759	8.77
	White	2,870	33.17
	More than one race	416	4.81
	Unreported race	79	0.92
Citizenship Status	Non-U.S. citizen	677	7.82
	U.S. citizen	7,975	92.18
College/School	Arts & Sciences	4,561	52.72

TABLE 1			
Frequency Distribution of the Sample (Categorical Variables)			
Variable	Levels	Frequency	Percent
	Business	1,579	18.25
	Education	420	4.85
	Nursing	835	9.65
	Policy Studies	291	3.36
	Undeclared major	966	11.17
Campus Residency	Did not live on campus	3,461	40.00
	Lived on campus	5,191	60.00
AP Credit Transfer	Did not transfer AP credits	5,906	68.26
	Transferred AP credits	2,746	31.74
FLC Participation	Did not participate in FLC	4,025	46.52
	Participated in FLC	4,627	53.48
Unmet Need Quartile	Bottom quartile	2,536	29.86
	2nd quartile	1,711	20.15
	3rd quartile	2,122	24.99
	Top quartile	2,124	25.01
Student Cohort	Fall 2010 cohort	2,831	32.72
	Fall 2011 cohort	2,699	31.20
	Fall 2012 cohort	3,122	36.08

TABLE 2				
Descriptive Statistics on Continuous/Discrete Variables				
	Mean	SD	Min	Max	N
First-term GPA	3.00	0.87	0.00	4.30	8,652
Age at Matriculation	18.48	0.60	16.30	49.80	8,652
High School GPA	3.36	0.32	2.10	4.00	8,618
SAT Verbal Score	541.60	71.39	280.00	800.00	8,410
SAT Math Score	544.90	73.92	250.00	800.00	8,413
Credit Load	14.34	1.387	12.00	21.00	8,652

given library resource more than their male counterparts do and that these female students also tend to have a higher term GPA compared to their male counterparts. In this case, a student input (sex) is related to both the environment (library resource utilization) and the output (term GPA). As Astin indicated, in such a case, the relationship between the environment and the output may simply reflect the effect of the input rather than an actual effect of the environment on the output.[23] In our example, this means that the effect of the input (sex) will be incorporated in the estimated effect of the environment (library use), thus causing the effect of the environment to be biased upward or downward.

To deal with this issue, we used propensity score matching. This approach was introduced by Rosenbaum and Rubin.[24] The propensity score is defined as "the conditional probability of exposure to the treatment given the observed covariates."[25] The propensity score makes it possible to deal with a critical empirical issue: it allows estimating effects of certain groups when random assignment is not possible and when individuals have "self-selected themselves into treatment or control conditions."[26] This approach has been used in higher education research to study different topics. For example, Attewell and his colleagues used it to investigate the impact of taking remedial courses on graduation and time to degree.[27] Schudde used it to examine the effect of campus residency on student retention.[28] Most recently, Chiteng Kot used it to estimate the impact of academic advising on first-year GPA and attrition.[29]

We used propensity score matching as a data preprocessing step.[30] First, we used logistic regression to predict students' use of library resources. We used each of the six treatment indicators of library resource utilization as the dependent variables and student characteristics (demographics, academic preparation, and campus experiences) as independent variables. For each student we generated the propensity score: in other words, the probability of using a particular library resource. We used these propensity scores to match individuals who used a particular library resource with those who did not use that recourse, such that the two groups had similar or almost identical background characteristics. As an example, we matched students who used workstations at least once with students who had a similar propensity score but who never used any workstations. We repeated this process for each of the remaining treatment indicators (using workstations at least five times, at least 10 times, and at least 20 times; using study rooms at least once; and attending research clinics at least once). The propensity score is particularly attractive because of its balancing property.[31] Rosenbaum and Rubin demonstrated that treated and control subjects with the same propensity score have the same distribution relative to the observed covariates.[32]

In this study, we used a matching approach known as nearest neighbor.[33] This approach consists in matching or pairing each person with a given propensity score in the treatment group with a person with the closest propensity score in the control group. We used matching with replacement. According to Stuart's review of matching methods, matching with replacement can decrease bias because control units that look similar to many treated units can be used multiple times.[34] Thus, we assigned each individual in the treatment group to one individual in the control group. Some individuals in the control group, however, were also assigned to multiple individuals in the treatment group (depending on how close the propensity scores were). In the final estimation of the treatment effect, we used frequency weights to adjust for the fact that some matched control units were used more than once. We discarded from our postmatching analyses individuals whose propensity score did not fall under the area of common support—in other words, the area where the ranges of propensity scores between treated and control cases overlap. Excluding these individuals from postmatching was essential to only compare groups that were similar.[35]

After matching, one needs to assess the covariate balance, or the similarity of the distributions of the set of covariates, between the treatment and control groups. For each covariate included in the analysis, we computed the *standardized difference in percent*: that is, the "mean difference as a percentage of the average standard deviation."[36] Based on Rosenbaum and Rubin's measure, an absolute standardized difference in percent of 20 is often used as a threshold.[37] This implies that the difference between treatment and control groups should be no more than 20 percent of a standard deviation.

Results

Logistic Regression Results

Table 3 presents the results of the logistic regression analysis of using library resources. Results indicate that the probability of using library resources was a function of student characteristics: clearly, some students were more likely to use library resources than other students. Asian students were more likely to use workstations compared to Black students, when workstation usage was defined as (1) having logged in at least once and (2) having logged in at least five times. For example, when workstation usage was defined as having logged in to a workstation at least five times during the term, Asian students were found to be 31 percent more likely to use workstations than Black students. Hispanic students and White students, in contrast, were less likely to use workstations compared to Black students. For example, when workstation usage was defined as having logged in to workstations at least 20 times during the term, Hispanic students were found to be 32 percent less likely and White students 57 percent less likely to use workstations than Black students. Regardless of how workstation usage was defined, non-U.S. citizens were consistently more likely to use workstations compared to U.S. citizens. Likewise, students who lived on campus and those who participated in FLCs were less likely to use workstations, compared to their counterparts who did not live on campus or participate in FLCs. Students in higher quartiles of unmet financial need were more likely to use workstations compared to their counterparts in the lowest quartile (students who did not have any unmet need). The student's credit load also appeared to be positively related to the likelihood of using library workstations. For instance, when the number of credits that the students took increased by 1, the likelihood of using workstations at least once increased by 8 percent, and the likelihood of using workstations at least 20 times increased by 10 percent.

With respect to study room use, female students were about 29 percent more likely to use study rooms than male students. Asian students were 253 percent more likely and Hispanic students 49 percent more likely to use study rooms than Black students. Students who transferred AP credits at matriculation were 39 percent more likely to use study rooms than their counterparts who did not transfer any AP credits. Students who participated in FLCs were 38 percent more likely to use study rooms than those who did not participate in FLCs. The likelihood of using study rooms decreased with the student's age at matriculation but increased with the student's high school GPA. For instance, when the student's high school GPA increased by one point, the likelihood of using study rooms increased by 43 percent. Finally, with respect to research clinic attendance, students who participated in FLCs were 705 percent (a staggering difference) more likely to attend research clinics than their counterparts who did not participate in FLCs. Also, the likelihood of attending research clinics decreased slightly with a student's SAT math score.

In summary, logistic regression results in table 3 suggest that students who used a particular library resource at the level specified differed from those who did not use that resource at the level specified. Therefore, we used logistic regression results to generate predicted probabilities of using library resources at each of the levels specified. We then used these predicted probabilities, which were estimated propensity scores, to match students in the treatment groups (students who used a given library resource) with those in the control groups (students who did not use that library resource). We discarded from the analyses students whose propensity scores did not fall under the area of common support. One average, 33 students had propensity scores that fell outside the area of common support.

	Used Workstations at Least 1 Time		Used Workstations at Least 5 Times		Used Workstations at Least 10 Times	
TABLE 3 (PART 1) **Predictors of Library Resource Utilization**						
	Odds Ratio	SE	Odds Ratio	SE	Odds Ratio	SE
Female	1.034	0.057	0.97	0.051	0.933	0.055
Male (Reference)						
Asian	1.332**	0.125	1.310***	0.102	1.164	0.097
Hispanic	0.886	0.088	0.757**	0.068	0.811*	0.081
White	0.545***	0.036	0.493***	0.033	0.490***	0.038
More than One Race	0.966	0.12	0.825	0.094	0.866	0.112
Unreported Race	1.098	0.356	1.117	0.307	1.353	0.387
Black (Reference)						
Non-U.S. Citizen	1.655***	0.206	1.449***	0.141	1.425***	0.142
U.S. Citizen (Reference)						
Age at Matriculation	0.965	0.04	0.973	0.038	0.974	0.041
High School GPA	1.058	0.087	1.123	0.088	1.217*	0.108
SAT Verbal Score	1.007	0.004	1.001	0.004	0.997	0.004
SAT Math Score	0.996	0.004	0.993	0.004	1	0.004
AP Credit Transfer	1.009	0.059	1.119*	0.062	1.079	0.067
No AP Credit Transfer (Reference)						
Business	0.961	0.066	0.918	0.061	0.977	0.073
Education	0.927	0.108	0.805	0.095	0.806	0.112
Nursing	1.179	0.111	1.132	0.094	1.198*	0.109
Policy Studies	0.978	0.137	0.957	0.129	1.021	0.156
Undeclared Major	1.015	0.085	0.941	0.075	0.979	0.088
Arts and Sciences (Reference)						
Credit Load	1.080***	0.021	1.046*	0.019	1.035	0.021
Campus resident	0.637***	0.035	0.505***	0.026	0.449***	0.026
Non-campus Resident (Reference)						
FLC Participant	0.934	0.052	0.849**	0.045	0.830**	0.049
No FLC Participant (Reference)						
Unmet Need: 2nd Quartile	1.376***	0.098	1.350***	0.096	1.328***	0.111
Unmet Need: 3rd Quartile	1.539***	0.107	1.506***	0.103	1.605***	0.127
Unmet Need: Top Quartile	1.734***	0.126	1.625***	0.112	1.676***	0.133
Unmet Need: Bottom Quartile (Reference)						

TABLE 3 (PART 1) Predictors of Library Resource Utilization					
	Used Workstations at Least 1 Time		Used Workstations at Least 5 Times		Used Workstations at Least 10 Times
	Odds Ratio	SE	Odds Ratio	SE	Odds Ratio	SE
Fall 2011 Cohort	1.049	0.065	0.908	0.055	0.873*	0.06
Fall 2012 Cohort	1.164*	0.072	1.087	0.064	1.1	0.072
Fall 2013 Cohort (Reference)						
Chi-square (df)	514.00 (25)		742.49 (25)		654.86 (25)	
N	8232		8232		8232	

TABLE 3 (PART 2) Predictors of Library Resource Utilization					
	Used Workstations at Least 20 Times		Used Study Rooms at Least 1 Time		Attended at Least 1 Research Clinic
	Odds Ratio	SE	Odds Ratio	SE	Odds Ratio	SE
Female	0.934	0.075	1.287*	0.128	1.307	0.214
Male (reference)						
Asian	1.072	0.116	3.535***	0.492	1.328	0.348
Hispanic	0.680**	0.093	1.492*	0.255	1.091	0.282
White	0.433***	0.049	1.261	0.166	1.428	0.278
More than One Race	0.715	0.134	1.102	0.255	1.574	0.47
Unreported Race	1.027	0.389	5.466***	2.272	1.532	1.693
Black (Reference)						
Non-U.S. Citizen	1.469**	0.177	0.971	0.161	0.92	0.331
U.S. Citizen (Reference)						
Age at Matriculation	0.984	0.052	0.715***	0.072	1.108	0.178
High School GPA	1.509***	0.181	1.432*	0.212	1.077	0.258
SAT Verbal Score	1	0.006	0.980**	0.007	0.981	0.012
SAT Math Score	1.002	0.006	1.001	0.007	0.965**	0.012
AP Credit Transfer	1.004	0.085	1.388**	0.141	0.871	0.174
No AP Credit Transfer (Reference)						
Business	0.85	0.088	0.982	0.121	0.705	0.15
Education	0.698	0.143	0.976	0.211	1.187	0.352
Nursing	1.055	0.128	0.974	0.148	1.193	0.277
Policy Studies	0.982	0.208	1.038	0.265	0.832	0.329
Undeclared Major	0.889	0.108	0.791	0.127	1.251	0.298
Arts and Sciences (Reference)						

	Used Workstations at Least 20 Times		Used Study Rooms at Least 1 Time		Attended at Least 1 Research Clinic	
	Odds Ratio	SE	Odds Ratio	SE	Odds Ratio	SE
Credit Load	1.100***	0.03	1.022	0.035	1.017	0.058
Campus Resident	0.324***	0.026	0.953	0.092	0.939	0.152
Non-campus Resident (Reference)						
FLC Participant	0.762***	0.062	1.383**	0.141	8.053***	1.725
No FLC Participant (Reference)						
Unmet Need: 2nd Quartile	1.267*	0.146	0.787	0.114	1.1	0.266
Unmet Need: 3rd Quartile	1.442***	0.159	0.826	0.109	1.278	0.272
Unmet Need: top Quartile	1.500***	0.164	1.127	0.139	1.426	0.307
Unmet Need: Bottom Quartile (Reference)						
Fall 2011 Cohort	0.699***	0.067	—	—	—	—
Fall 2012 Cohort	1.086	0.094	0.763**	0.07	—	—
Fall 2013 Cohort (Reference)						
Chi-square (df)	516.1 (25)		199.84 (24)		203.74 (23)	
N	8232		5524		2954	

TABLE 3 (PART 2)
Predictors of Library Resource Utilization

Results of Propensity Score Matching

After matching individuals in the treatment group with those in the control group, one needs to assess balance to ensure that treatment and control groups are similar. Table 4 presents a summary of covariate balance both before and after matching. This table identifies the number of predictors at different levels of the standardized bias, which was expressed in terms of the absolute *standardized difference in percent*. For example, when the treatment was defined as using workstations at least once during the semester, four of the predictors had an absolute standard difference in percent that was 25 or greater. This means that students who used workstations at least once during the term differed from those who did not use workstations by at least a quarter of a standard deviation. However, after matching, the absolute standardized difference in percent for each of the predictors of using workstations at least once fell below 10.

Across the six treatment indicators, all the predictors had a standardized difference less than 10, after matching, with the exception of four predictors of research clinic attendance, which had values between 10 and 14.9. On average, after matching, more than eight in ten predictors had a standardized difference that was less than 5 percent; less than two in ten predictors had a standardized difference equal to or greater than 5 percent (but also less than 15 percent). These values fell well below the 20 percent threshold that is often used based on Rosenbaum and Rubin's measure.[38]

TABLE 4						
Summary of Standard Bias before and after Matching: Number of Predictors at Each Level of the Standard Bias						
	Used Workstations at Least 1 Time		Used Workstations at Least 5 Times		Used Workstations at Least 10 Times	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
Standard Bias (in %)						
25 or greater	4		4		4	
15 to 24.9	2		2		4	
10 to 14.9	3		3			
5 to 9.9	7	1	5	2	8	1
Less than 5	13	28	15	27	13	28
Standard Bias for the Propensity Score	55.52	0.00	63.01	0.00	65.65	0.00
	Used Workstations at Least 20 Times		Used Study Rooms at Least 1 Time		Attended at Least 1 Research Clinic	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
Standard Bias (in %)						
25 or greater	5		1		4	
15 to 24.9	3		5		2	
10 to 14.9	3		7		7	4
5 to 9.9	3	3	5	7	6	12
Less than 5	15	26	9	20	7	10
Standard Bias for the Propensity Score	76.45	0.02	57.35	0.01	99.00	0.04
Note: Standard bias is expressed in terms of the standardized difference in percent.						

Results from these analyses indicated that, before matching, the treatment and control groups differed significantly along many of the observed covariates. However, matching substantially decreased bias and made treatment and control subjects similar along the observed covariates. This approach allowed us to eliminate, or at least substantially decrease, the relationship between library resource utilization and student characteristics, which in turn made it possible to estimate the impact of library resource usage more accurately.

Results of Parametric Analyses after Matching

After matching, one can examine the mean difference in the outcome of interest—term GPA in the present analysis—between the treatment and control groups. However, some researchers have argued that this bivariate analysis could still result in bias. Thus, the use of parametric methods after matching has been

TABLE 5
Average Treatment Effect (ATE) from OLS Regression: Differences in GPA between Treatment and Control Groups after Matching

Group	OLS Regression		Sample Size	
	ATE	Cohen's d	Treatment Group	Control Group
Used Workstations at Least 1 Time	0.123***	0.15	5,753	1,835
Used Workstations at Least 5 Times	0.155***	0.19	3,234	1,925
Used Workstations at Least 10 Times	0.174***	0.21	2,021	1,511
Used Workstations at Least 20 Times	0.157***	0.20	921	791
Used Study Rooms at Least 1 Time	0.196***	0.26	593	529
Attended 1 or More Research Clinics	0.183*	0.24	227	204

***: $p < 0.001$; **: $p < 0.01$; * $p < 0.05$
The relative size of Cohen's *d* values indicates a *negligible* effect when *d* is < 0.15; a *small* effect when *d* is >= 0.15 and < 0.40; a *medium* effect when *d* is >= 0.40 and < 0.75, and a *large* effect when *d* is >=0.75.

TABLE 6 (PART 1)
Results of OLS Regression Analysis for Academic Performance after Matching

	Used Workstations at Least 1 Time		Used Workstations at Least 5 Times		Used Workstations at Least 10 Times	
	Coefficient	*Std. Error*	*Coefficient*	*Std. Error*	*Coefficient*	*Std. Error*
Treatment	0.123***	0.022	0.155***	0.023	0.174***	0.026
Female	0.059**	0.021	0.050*	0.025	0.020	0.029
Asian	0.173***	0.030	0.167***	0.035	0.174***	0.039
Hispanic	0.126***	0.035	0.146***	0.043	0.096*	0.049
White	0.098***	0.026	0.113***	0.034	0.163***	0.040
More than 1 Race	0.039	0.044	0.075	0.057	0.123*	0.061
Unreported Race	0.239*	0.109	0.248*	0.118	0.233	0.140
Non-U.S. Citizen	0.119**	0.037	0.094*	0.040	0.128**	0.046
Age at Matriculation	–0.026	0.016	–0.018	0.023	–0.031	0.026
High School GPA	0.874***	0.031	0.857***	0.037	0.826***	0.044
Sat Verbal Score	0.006***	0.002	0.006***	0.002	0.007**	0.002
Sat Math Score	0.001	0.002	0.001	0.002	0.001	0.002
AP Credit Transfer	0.188***	0.022	0.174***	0.026	0.227***	0.030
Business	–0.010	0.026	–0.014	0.032	–0.012	0.037
Education	0.015	0.046	0.021	0.059	0.063	0.071
Nursing	–0.031	0.032	–0.018	0.039	–0.019	0.043
Policy Studies	0.064	0.054	0.011	0.067	0.012	0.079
Undeclared Major	0.027	0.031	0.035	0.037	0.042	0.044

TABLE 6 (PART 1)					
Results of OLS Regression Analysis for Academic Performance after Matching					
	Used Workstations at Least 1 Time		**Used Workstations at Least 5 Times**		**Used Workstations at Least 10 Times**	
	Coefficient	*Std. Error*	*Coefficient*	*Std. Error*	*Coefficient*	*Std. Error*
Credit Load	0.018**	0.007	0.024**	0.008	0.023*	0.010
On-campus Residency	0.106***	0.020	0.131***	0.024	0.097***	0.028
FLC Participation	0.075***	0.021	0.060*	0.025	0.103***	0.029
Unmet Need 2nd Quartile	−0.013	0.028	0.005	0.036	−0.004	0.043
Unmet Need 3rd Quartile	−0.050	0.027	−0.060	0.034	−0.050	0.040
Unmet Need 4th Quartile	−0.070**	0.027	−0.051	0.034	−0.064	0.040
Fall 2011 Cohort	−0.013	0.024	0.012	0.029	0.029	0.034
Fall 2012 Cohort	0.046*	0.023	0.066*	0.028	0.084**	0.032
Constant (Intercept)	−0.467	0.335	−0.662	0.477	−0.273	0.534
R-squared	0.155		0.152		0.164	
Adjusted R-squared	0.152		0.148		0.158	
F	53.142		35.290		26.438	
Degrees of Freedom	(26, 7560)		(26, 5131)		(26, 3504)	
N	7587		5158		3531	

TABLE 6 (PART 2)					
Results of OLS Regression Analysis for Academic Performance after Matching					
	Used Workstations at Least 20 Times		**Used Study Rooms at Least 1 Time**		**Attended at Least 1 Research Clinic**	
	Coefficient	*Std. Error*	*Coefficient*	*Std. Error*	*Coefficient*	*Std. Error*
Treatment	0.157***	0.037	0.196***	0.045	0.183*	0.073
Female	0.003	0.041	0.031	0.051	0.136	0.083
Asian	0.149**	0.054	0.180*	0.070	0.152	0.136
Hispanic	0.202**	0.072	0.063	0.087	0.189	0.138
White	0.102	0.060	0.119	0.069	−0.141	0.098
More than 1 Race	0.014	0.102	0.102	0.118	0.126	0.144
Unreported Race	0.516*	0.202	0.336	0.222	−0.652	0.784
Non-U.S. Citizen	0.137*	0.060	0.033	0.083	0.279	0.212
Age at Matriculation	−0.068	0.035	−0.161**	0.051	−0.031	0.089
High School GPA	0.795***	0.062	0.925***	0.075	0.914***	0.128

	Used Workstations at Least 20 Times		Used Study Rooms at Least 1 Time		Attended at Least 1 Research Clinic	
	Coefficient	*Std. Error*	*Coefficient*	*Std. Error*	*Coefficient*	*Std. Error*
Sat Verbal Score	0.007*	0.003	0.004	0.004	0.011	0.007
Sat Math Score	0.001	0.003	–0.006	0.004	0.002	0.007
AP Credit Transfer	0.203***	0.044	0.207***	0.051	0.088	0.110
Business	–0.074	0.053	0.052	0.065	0.003	0.108
Education	0.088	0.109	0.248*	0.108	–0.056	0.148
Nursing	–0.003	0.061	–0.008	0.074	0.026	0.122
Policy Studies	–0.074	0.111	0.222	0.141	–0.234	0.231
Undeclared Major	–0.047	0.061	0.084	0.081	–0.049	0.128
Credit Load	0.021	0.014	–0.011	0.017	–0.042	0.032
On-campus Residency	0.100*	0.041	0.088	0.048	0.198*	0.083
FLC Participation	0.055	0.042	0.131*	0.053	–0.042	0.116
Unmet Need 2nd Quartile	0.026	0.061	0.056	0.074	0.017	0.124
Unmet Need 3rd Quartile	0.011	0.058	–0.052	0.068	–0.202	0.106
Unmet Need 4th Quartile	–0.031	0.057	–0.011	0.063	–0.302**	0.109
Fall 2011 Cohort	–0.004	0.050				
Fall 2012 Cohort	0.061	0.045	0.138**	0.047		
Constant (Intercept)	0.571	0.750	2.696**	1.024	0.326	1.779
R-Squared	0.162		0.182		0.199	
Adjusted R-squared	0.150		0.163		0.151	
F	12.561		9.758		4.198	
Degrees of Freedom	(26, 1684)		(25, 1096)		(24, 406)	
N	1711		1122		431	

TABLE 6 (PART 2)
Results of OLS Regression Analysis for Academic Performance after Matching

Note: The header row indicates how the treatment indicator was defined.

suggested to avoid omitted variable bias, adjust for any covariate imbalance remaining after matching, and estimate the treatment effect based on a model that is robust against misspecification.[39] Table 5 summarizes average treatment effects (ATEs) estimated after adjusting for student demographic characteristics, academic preparation, and campus experiences. Table 6 presents full regression results using matched samples. As results indicate, all the ATEs were statistically significant, implying that using library resources at the levels specified was associated with higher first-term GPA.

FIGURE 2
Mean GPA after Adjusting for Student Characteristics
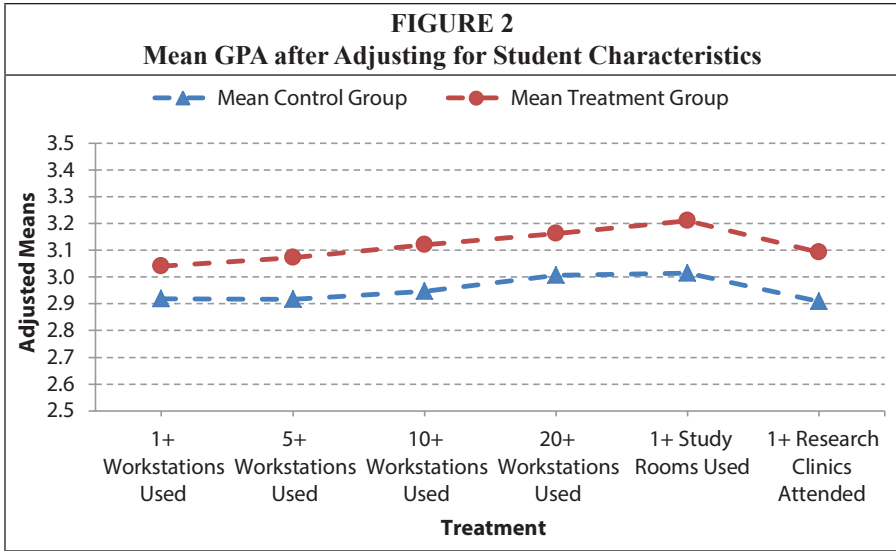
The gain in term GPA was between 12.3 percentage points and 19.6 percentage points across the six treatment indicators. Results indicate that using workstations at least once as opposed to not using them at all during the term was associated with a gain of 12 percentage points in term GPA. This gain was 15.5 percentage points for students who used workstations at least 5 times (compared to those who used them 0 to 4 times), 17.4 percentage points for students who used workstations at least 10 times (compared to those who used them 0 to 9 times), and 15.7 percentage points for those who used workstations at least 20 times (compared to their counterparts who used them 0 to 19 times). Study room usage and research clinic attendance were associated with a gain of 19.6 and 18.3 percentage points respectively for students who used these resources compared to their counterparts who did not use them.

To assess whether these gains were of practical significance, we computed the effect size (*Cohen's d*) for each ATE. *Cohen's d* values varied between 0.15 and 0.26. Each value suggested that the corresponding treatment effect was small but not negligible. Figure 2 provides an illustration of adjusted means, after controlling for student characteristics. For each treatment indicator, the adjusted mean for students in the treatment groups was always higher compared to the adjusted mean for their counterparts in the control groups.

Summary and Discussion

This study sought to examine the relationship between library resource utilization defined as workstation usage, study room usage, and research clinic attendance and undergraduate academic performance measured by first-term GPA. To reduce self-selection bias, we used propensity score matching to construct treatment and control groups that were similar on background characteristics. We found that using library resources at each of the levels specified was associated with a small but also meaningful gain in first-term GPA. With respect to the magnitude of the effects, the largest gain in term GPA was associated with using study rooms at least once during the term (gain of 20 percentage points on term GPA) and attending research clinics at least once (18 percentage points). The effect sizes corresponding to these two "treatment effects" suggest that the mean term GPA of students in the treatment group was about a quarter of a standard deviation above the mean GPA for students in the control group. An effect size conversion

methodology developed by Coe provides perhaps a more practical interpretation.[40] For study room usage, an effect size of 0.26 implies that the term GPA for the average student in the treatment group who used study rooms at least once during the term exceeded the term GPA of around 62 percent of the students in the control group who never used study rooms. For research clinic attendance, the effect size of 0.24 implies that the term GPA for the average student in the treatment group who attended at least one research clinic during the term exceeded the term GPA of around 58 percent of the students in the control group, those students who never attended a research clinic.

With workstation usage defined as using library workstations at least five times (compared to 0 to 4 times), at least 10 times (compared to 0 to 9 times), or at least 20 times (compared to 0 to 19 times), the effect sizes suggest that the mean term GPA of students in the treatment groups was about one-fifth of a standard deviation above the mean GPA for students in the control group. This result further suggests that the term GPA for the average student in the treatment group who used workstations at the levels specified exceeded the term GPA of around 58 percent of the students in the corresponding control group. When workstation usage was defined as using workstations at least once (compared to zero times), results suggested that the mean term GPA of students who used workstations at least once was about 15 percent of a standard deviation above the mean GPA of students who never used a library workstation. This last finding further suggested that the term GPA for the average student who used workstations at least once during the first term exceeded the term GPA of around 54 percent of the students who never used any workstations during the term.

Thus, using library resources as defined in this research project was positively related with first-term GPA. Regardless of the threshold values used to define library resource utilization, the gain in first-term GPA appeared to be meaningful, albeit small. Because of the substantial difference in the methodological approach, statistical results from the present study cannot be directly compared with those of previous studies. In substance, however, findings from this study support the notion that there is a positive relationship between students' use of library resources and academic outcomes. The quasi-experimental research design used in this study allows us to conclude that the academic performance of an average student who used a given library resource during his/her first term was higher than that of most students who did not use that resource during their first term.

This study now provides evidence that the library has an impact on the academic performance of first-time, full-time undergraduate students at Georgia State University. The results of this study provide relevant campus units, the library in particular, with "a compelling story to share based on the data."[41] While the library already encourages faculty to urge students to seek out the library's services and resources, findings from this study may help convey to faculty that there is a proven, positive relationship between library usage and students' academic performance (GPA): Students who use the library tend to have a higher GPA compared to those who do not. The library can use this information in a marketing campaign to students, communicating to them that a known characteristic of a successful student is that of library user. This evidence also can be relayed to other student support units, such as academic advising and first-year programs, to aid them in guiding and making recommendations to struggling students. The implication for campus administration is that the library is not just a passive study space; it contributes to student success and, consequently, to the pursuit of institutional goals and objectives. Thus, investing in the university library is investing in student success. Library employees who hope that their efforts make a difference in regard to students' academic performance now know that the services and resources they provide contribute positively to student success. This knowledge should help library

employees in outreach to non–library users and also in their interactions with library users, by assuring them that using library resources is a path toward academic success.

Limitations and Future Research

This study has a number of limitations, some of which are relative to the analytical procedures and others to the data used. With respect to the analytical procedures, it should be noted that propensity score matching seeks to achieve balance on observed covariates only.[42] Some researchers have shown that propensity score matching may also help reduce selection bias due to unobserved covariates.[43] However, it may still be possible that this analysis has not included an important variable that predicts both library resource utilization and first-term GPA, which may lead to bias in the estimation of the average treatment effect. An example of a factor that is difficult to observe or control for is motivation. If students who use library resources are more motivated than those who do not use library resources, and if these motivated students also earn higher grades than less motivated students, then the estimated treatment effects may be biased. Findings from this study should therefore be interpreted cautiously and should be viewed as reflecting a relationship, not a cause and effect.

Several limitations can also be noted relative to the data used to measure library resource utilization. First, it is unclear what students actually do when they log in to library workstations. Although one can assume that students use library workstations for academic purposes (such as research, homework, and online learning activities), it may also be possible that many students use library workstations for nonacademic purposes. This makes it difficult to understand how the use of library workstations actually impacts term GPA. Second, with respect to study room usage, although the library requires online reservations for study rooms, it is possible that some students simply walk in and use an available study room without a reservation. Another important consideration is that when a student reserves a room for group use, the online reservation system creates a record for that student only, and no records are created for other group members. Thus, students who never used their campus username to reserve a study room but who actually used a study room were not classified as having used any study rooms. Finally, the library provides various important resources and services that were not captured in this research project due to a lack of historical data. These resources and services include, among other things, loans and access to electronic resources. Thus, the resources included in this research project provide only a limited picture of how library resource utilization impacts student academic performance.

Georgia State University Library has begun to collect data on other resources, such as off-campus logins to electronic resources, and is investigating the possibility of collecting other data types, including material loans, interlibrary loan requests, and library instruction attendance. In the future, the research design used in the present study will be expanded to other library resources and services. In addition, the library and the Office of Institutional Effectiveness plan to track students' use of library resources over time, beginning with students' first term on campus. This longitudinal dataset will allow for an in-depth analysis of the impact of library resource usage on various academic outcomes, including GPA, retention, and graduation.

Conclusion

In an era of accountability in which higher education institutions and campus units are increasingly asked to demonstrate their value, assessing the impact of academic resources and services on student success has become perhaps more critical than ever. Traditionally, the academic library has been a vital part of a student's academic life. Academic libraries are uniquely situated in that they can be modified to offer college students a better aca-

demic experience. Through the resources and services that they offer, academic libraries can create or foster an environment that is conducive to student learning and success. Unfortunately, research on the contribution of academic libraries to student success has lagged behind and is still in its infancy. The present study controlled for student inputs, environmental experiences on campus, and self-selection and measured—in a quantifiable way—the positive impact of library resource utilization on the academic success of new college students. The library, as this study shows, has a positive impact on students' academic performance, net of other factors (student demographic characteristics, precollege academic preparation, and other environmental experiences on campus)." The present study particularly contributes to the body of research on the impact of academic libraries by (1) examining this impact through the lens of Alexander Astin's Input-Environment-Output conceptual framework, (2) using analytical procedures that account for self-selection bias, and (3) estimating the average treatment effect and its corresponding effect size. It is our hope that this study will inspire researchers to use similar analytical tools to further investigate the impact of academic libraries on student success.

Notes

1. Krista M. Soria, Jan Fransen, and Shane Nackerud, "Library Use and Undergraduate Student Outcomes: New Evidence for Students' Retention and Academic Success," *portal: Libraries and the Academy* 13, no. 2 (2013): 147.

2. Association of College and Research Libraries, *The Value of Academic Libraries: A Comprehensive Research Review and Report,* res. Megan Oakleaf (Chicago: Association of College and Research Libraries, 2010), 5, available online at www.ala.org/acrl/sites/ala.org.acrl/files/content/issues/value/val_report.pdf [accessed 7 November 2013].

3. Brian L. Cox and Margie Jantti, "Capturing Business Intelligence Required for Targeted Marketing, Demonstrating Value, and Driving Process Improvement," *Library & Information Science Research* 34, no. 4 (2012): 308–16.

4. Timothy Renick, "Education Advisory Board/Complete College Georgia Update" (presentation, Georgia State University Library, Atlanta, Ga., September 27, 2012.)

5. Ernest T. Pascarella and Patrick T. Terenzini, *How College Affects Students: A Third Decade of Research* (San Francisco: Jossey-Bass, 2005), 396.

6. Shun Han Rebekah Wong and T.D. Webb, "Uncovering Meaningful Correlations between Student Academic Performance and Library Material Usage," *College & Research Libraries* 72, no. 4 (2011): 362.

7. Association of College and Research Libraries, *The Value of Academic Libraries*, 1–182.

8. Ibid., 14, 178.

9. Wong and Webb, "Uncovering Meaningful Correlations between Student Academic Performance and Library Material Usage," 361–70.

10. Shun Han Rebekah Wong and Dianne Cmor, "Measuring Association between Library Instruction and Graduation GPA," *College & Research Libraries* 72, no. 5 (2011): 464–73.

11. Cox and Jantti, "Capturing Business Intelligence," 1.

12. Cox and Jantti, "Capturing Business Intelligence," 308–16; Brian L. Cox and Margie Jantti, "Discovering the Impact of Library Use and Student Performance," *EDUCAUSE Review Online* (2012), available online at www.educause.edu/ero/article/discovering-impact-library-use-and-student-performance [accessed 11 November 2013].

13. Cox and Jantti, "Capturing Business Intelligence," 311.

14. Deborah Goodall and David Pattern, "Academic Library Non/Low Use and Undergraduate Student Achievement," *Library Management* 32, no. 3 (2011): 159–70.

15. Graham Stone and Byrony Ramsden, "Library Impact Data Project: Looking for the Link between Library Usage and Student Attainment," *College & Research Libraries* 74, no. 6 (2013): 546–59.

16. Ibid., 554.

17. Ed Cherry, Stephanie Havron Rollins, and Toner Evans, "Proving Our Worth: The Impact of Electronic Resource Usage on Academic Achievement," *College & Undergraduate Libraries* 20, no. 3–4 (2013): 386–98.

18. Soria, Fransen, and Nackerud, "Library Use and Undergraduate Student Outcomes," 151.

19. Ibid., 154.

20. Ibid.

21. Alexander W. Astin, "The Methodology of Research on College Impact, Part One," *Sociology of Education* 43, no. 3 (1970): 223–54; Alexander W. Astin, "The Methodology of Research on College Impact, Part Two," *Sociology of Education* 43, no. 4 (1970): 437–50; Alexander W. Astin, *Assessment for Excellence: The Philosophy and Practice of Assessment and Evaluation in Higher Education* (New York: MacMillan, 1991).

22. Astin, "The Methodology of Research, Part One," 64.

23. Astin, *Assessment for Excellence*.

24. Paul R. Rosenbaum and Donald B. Rubin, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika* 70, no. 1 (1983): 41–55.

25. Paul R. Rosenbaum, *Design of Observational Studies* (Philadelphia: Springer, 2010).

26. Barbara Schneider, Martin Carnoy, Jeremy Kilpatrick, and Richard J. Shavelson, *Estimating Causal Effects Using Experimental and Observational Designs* (Washington, D.C.: American Educational Research Association, 2007), 50.

27. Paul A. Attewell, "New Evidence on College Remediation," *Journal of Higher Education* 77, no. 5 (2006): 886–924.

28. Lauren T. Schudde, "The Causal Effect of Campus Residency on College Student Retention," *Review of Higher Education* 34, no. 4 (2011): 581–610.

29. Felly Chiteng Kot, "The Impact of Centralized Advising on First-Year Academic Performance and Second-Year Enrollment Behavior," *Research in Higher Education* (2014), doi: 10.1007/s11162-013-9325-4.

30. Daniel Ho, Kosuke Imai, Gary King, and Elizabeth Stuart, "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference," *Political Analysis* 15, no. 3 (2007): 199–236.

31. Rosenbaum and Rubin, "The Central Role of the Propensity Score."

32. Ibid.

33. Donald B. Rubin, "Matching to Remove Bias in Observational Studies," *Biometrics* 29, no. 1 (1973): 159–83.

34. Elizabeth A. Stuart, "Matching Methods for Causal Inference: A Review and a Look Forward," *Statistical Science* 25, no. 1 (2010): 1–21.

35. Stephen L. Morgan and David J. Harding, "Matching Estimators of Causal Effects: Prospects and Pitfalls in Theory and Practice," *Sociological Methods & Research* 35, no. 1 (2006): 3–60; Ho, Imai, King, and Stuart, "Matching as Nonparametric Preprocessing"; Daniel Ho, Kosuke Imai, Gary King, and Elizabeth Stuart, "MatchIt: Nonparametric Preprocessing for Parametric Causal Inference," *Journal of Statistical Software* 42, no. 8 (2011): 1–28; Marco Caliendo and Sabine Kopeinig, "Some Practical Guidance for the Implementation of Propensity Score Matching," *Institute for the Study of Labor Discussion Paper No. 1588*, 2005, available online at http://ftp.iza.org/dp1588.pdf [accessed 24 October 2013].

36. Paul R. Rosenbaum and Donald B. Rubin, "Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score," *American Statistician* 39, no. 1 (1985): 34.

37. Lockwood Reynolds and Stephen L. DesJardins, "The Use of Matching Methods in Higher Education Research: Answering Whether Attendance at a 2-Year Institution Results in Differences in Education Attainment," in *Higher Education: Handbook of Theory and Research* 2009, ed. John C. Smart (Dordrecht, The Netherlands: Springer), 47–104.

38. Ibid.

39. Ho, Imai, King, and Stuart, "Matching as Nonparametric Preprocessing"; Kosuke Imai, Gary King, and Olivia Lau, "Toward a Common Framework for Statistical Analysis and Development," *Journal of Computational and Graphical Statistics* 17, no. 4 (2008): 892–913; Guido W. Imbens, "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review," *Review of Economics and Statistics* 86, no. 1 (2004): 4–29.

40. Robert Coe, "It's the Effect Size, Stupid: What Effect Size Is and Why It Is Important" (presentation at the Annual Conference of the British Educational Research Association, University of Exeter, England, Sept. 12–14, 2002), available online at www.leeds.ac.uk/educol/documents/00002182.htm [accessed 15 October 2013].

41. Cox and Jantti, "Discovering the Impact."

42. Donald B. Rubin, "Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation," *Health Services & Outcomes Research Methodology* 2 (2002): 169–88.

43. Thomas A. DiPrete and Henriette Engelhardt, "Estimating Causal Effects with Matching Methods in the Presence and Absence of Bias Cancellation," *Max Planck Institute for Demographic Research Working Paper 2000–013* (2000), available online at www.demogr.mpg.de/Papers/Working/wp-2000-013.pdf [accessed 11 November 2013].