Learning Communities Research and Practice

Volume 2 | Issue 1 Article 3

6-4-2014

Identical Profiles, Different Paths: Addressing Self-selection Bias in Learning Community Cohorts

Stephanie Zobac
Northern Illinois University, srzobac@noctrl.edu

Julia Spears
Northern Illinois University, spearsjulia@gmail.com

Gregory Barker
Northern Illinois University, gbarker1@niu.edu

Recommended Citation
Zobac, S. , Spears, J. , Barker, G. (2014). Identical Profiles, Different Paths: Addressing Self-selection Bias in Learning Community Cohorts. *Learning Communities Research and Practice*, 2(1), Article 3.
Available at: https://washingtoncenter.evergreen.edu/lcrpjournal/vol2/iss1/3

Identical Profiles, Different Paths: Addressing Self-selection Bias in Learning Community Cohorts

Abstract
This article presents a method for addressing the self-selection bias of students who participate in learning communities (LCs). More specifically, this research utilizes equivalent comparison groups based on selected incoming characteristics of students, known as bootstraps, to account for self-selection bias. To address the differences in academic preparedness in the fall 2012 cohort, three stratified random samples of students were drawn from the non-LC population to match the LC cohort in mean ACT composite scores and mean high school percentile ranks. This process is called bootstrapping. The study suggests that LCs do impact student academic achievement and retention. The results indicate that LC students with similar entering characteristics to those of the bootstrap sample had higher rates for both GPA and retention than non-LC participants.

Keywords
learning communities, self-selection, research methods, program assessment

Introduction

Learning communities (LCs) are a powerful national movement based on shifting pedagogy to focus on increasing integration across courses and on extending student learning beyond the classroom. One of the High-Impact Practices presented by George Kuh and the Association of American Colleges and Universities (2008), LCs consist of a cohort of students who co-enroll in two or more courses that often explore a common theme. By deliberately co-enrolling students, bundling courses, and developing a theme, LCs encourage students to integrate knowledge and learning across courses and disciplines while increasing interaction with peers and faculty (Gabelnick, MacGregor, Matthews, & Smith, 1990). Moreover, LCs often incorporate engaged learning opportunities that require students to apply their learning outside of the classroom through one or more of the following components: residential hall experiences, service-learning, field trips, and speakers (Kuh, 2008; Lenning & Ebbers, 1999; Smith, MacGregor, Matthews, & Gabelnick, 2004; Taylor, Moore, MacGregor, & Lindblad, 2004). In this way, LCs are purposefully designed to develop a sense of community, provide classroom or residential space to bring students together, create an engaging and supportive environment, require integration across academic and social experiences, develop interdisciplinary connections among LC courses, and provide the structure for students to develop higher order thinking (Brower & Dettinger, 1998). As LCs continue to gain momentum in higher education, it is not surprising that such practices also fall under increasing scrutiny.

One of the most common critiques of LCs is that of self-selection bias. Zhao and Kuh (2004) describe self-selection bias as the possibility "that students who choose to join a learning community are more academically able as reflected by measures of ability, which could account for differences in outcomes that might be associated with learning communities" (p. 120). Additional studies of LCs also indicate the potential issue of self-selection bias, including higher ACT scores (Pike, 1999; Inkelas & Weisman, 2003) and higher motivation (Stassen, 2003) of LC students. The tendency for more academically prepared and higher achieving students to enroll in LCs weakens the argument in favor of LCs. That is, any advantages LC students show at the conclusion of their participation in the LC might be attributed to the self-selection of more academically prepared students entering. As a result, LC administrators must determine if all students, including those with stronger admissions characteristics (e.g., ACT score, high school percentile rank, high school GPA), benefit from participating in LCs.

This article presents a method for addressing the self-selection bias in LCs. More specifically, this research utilizes equivalent comparison groups, known as bootstraps, to account for self-selection bias based on the above incoming characteristics of students to better determine the impact of LCs. The main

question guiding this research is: Does the bootstrap method adequately account for self-selection bias of LC students compared to non-LC students when measuring average GPA and retention?

Literature Review

Both theory and research drive the current learning community trend. Student development theory supports the need for programs like LCs that increase the academic and social involvement of students. Alexander Astin's (1993) student involvement theory indicates that the "effectiveness of any educational policy or practice is directly related to the capacity of that policy or practice to increase student involvement" (p. 519). Likewise, Vincent Tinto's (1993) departure theory reveals the importance of intellectual and social communities as students "make the transition to college and become incorporated or integrated into the life of the college" (p. 125). LCs support academic and social involvement by establishing communities that focus on a common theme (e.g., business ethics) and by providing space for students with similar interests to develop friendships.

In addition to foundational theories like Astin and Tinto, research indicates a multitude of benefits for students participating in LCs, ranging from increased academic success to openness to diversity. As a common outcome of university programming, research on LCs often investigates student academic success. Research indicates that participation in an LC leads to increased academic success, including higher GPAs (Huerta & Bray, 2013; Stassen, 2003; Tinto & Goodsell, 1993), increased number of credit hours earned (Baker & Pomerantz, 2000; Matthews, 1994; Tinto & Love, 1995), and increased academic effort (Zhao & Kuh, 2004). Research also suggests that students that participate in LCs are more likely to engage in habits that further foster academic success, such as higher levels of active and collaborative learning (Pike, Kuh, & McCormick, 2011; Smith, et al, 2004; Tinto, 1998). LC literature and research also indicates that LC participation leads to a smoother transition to college (Stassen, 2003) as well as increased retention and persistence (Brower & Inkelas, 2010; Stassen, 2003; Tinto & Russo, 1994).

In addition to academic success and increased retention and persistence, research suggests that LC participation has further benefits. Many studies suggest that LC participation results in higher levels of peer interaction (Cross, 1998; Inkelas & Weisman, 2003), faculty interaction (Cross, 1998; Inkelas, Vogt, Longerbeam, Owen, & Johnson, 2006), positive perception and overall satisfaction with college (Baker & Pomerantz, 2000; Brower & Inkelas, 2010), openness to diversity experiences (Pike et al, 2011; Tinto & Love, 1995), and integrative thinking (Matthews, 1994; Zhao & Kuh, 2004).

Purpose of This Study

The current article focuses on a learning communities program at a four-year public research university in the Midwest. In general, academic LCs consist of two to four "paired courses," which are individually taught courses linked by cohort enrollment and a theme that connects the course content (Shapiro & Levine, 1999). In the LC program, first-year students must co-enroll in the full LC bundle, which often includes a first-year seminar course. Currently, this program is limited to first-year students in their first semester. Faculty members are required to create at least two integrative assignments that draw on concepts from each course within the LC. A Peer Leader is assigned to each community to help LC students improve study habits and attendance rates and locate campus resources and to serves as a positive role model. Furthermore, LC students have the opportunity to engage in experiential learning opportunities such as speakers, field trips, and service learning that augment the theme or assignments of the LC, as well as program-wide events such as bowling that allow for socialization among LCs.

While research helps to guide the work of LC practitioners, it is often challenging for institutions to investigate the impact of their LC programs to the extent of the aforementioned research studies. It remains necessary for LC practitioners to measure the success of their LC programs in order to determine the impact and benefit for students, thus documenting the advantages of participating in LCs. Such validation allows LC administrators to better promote LCs on campus, increase support for LCs across campus, and contribute to the growing body of LC research.

In addition to the need for more research on the impact of LCs, it is also necessary for LC administrators to account for self-selection bias. Numerous articles list self-selection bias as a limitation within the study, which parallels practice as administration often indicates self-selection bias as a contributing factor in the positive findings associated with LC participation. The following method was designed to implement a quick LC assessment that accounts for self-selection bias, which is often the primary limitation of LC research and major concern of program coordinators and university administrators.

Methods

Participants

Table 1 presents the admissions characteristics of LC and non-LC students for the Fall 2010, 2011, and 2012 cohorts. Two observations are worth noting: First, the number of students participating in LCs has grown dramatically since

the inception of the program. The first official LCs ran in Fall 2010 when 46 students enrolled in two LCs. This number increased to 273 students enrolled in 13 LCs in Fall 2012. Second, the LC cohort has become increasingly more academically prepared than the non-LC cohort. Specifically, the Fall 2012 LC cohort has higher entering performance characteristics than the non-LC cohort in mean ACT composite scores, high school GPAs (HSGPA), and high school percentile ranks (HSCentile). As a result, any differences in success at the university (e.g., semester and cumulative GPA or retention) can be interpreted as the result of differences in academic preparedness (i.e., more well prepared students are more likely to join LCs and are more likely to succeed than less prepared students).

Table 1
Fall 2010, 2011, and 2012 LC and Non-LC Cohort Admissions Characteristics

	Fall 2010		Fall 2011		Fall 2012		
	Non-LC	LC	Non-LC	LC	Non-LC	LC	Bootstrap Samples
Frequency	2659	46	2340	250	2391	273	819
Mean ACT Composite	21.58	22.26	21.87	22.49	21.62	23.27	22.99
Mean HSGPA	3.07	3.04	3.05	3.19	3.09	3.35	3.26
Mean HSCentile	61.46	61.70	60.37	63.56	61.05	69.02	68.38

Academic indicators suggest LC students are more academically prepared when entering the university than non-LC students. Therefore, it is necessary to address the "possibility of self-selection (perhaps learning communities attract more academically capable students) or to the distinctive features of the learning community milieu that foster higher levels of student engagement" to determine the impact of LCs, if any (Zhao & Kuh, 2004, p. 123). This is especially true as the LC program continues to grow, with numerous LCs focusing on traditionally stronger student populations including honors, pre-medicine, engineering, nursing, and more.

Procedure

To address the differences in academic preparedness, a simple bootstrapping methodology was employed. Bootstrapping is a non-parametric approach to drawing statistical inferences by which a population is "resampled" multiple times to provide an estimate of the sampling distribution for a statistic (Mooney &

Duval, 1993). In the current study, three stratified random bootstrap samples of students were drawn (with replacement) from the 2012 non-LC population so that each sample matched the LC cohort in mean ACT Composite and mean HSCentile. Because HSGPA and HSCentile are highly correlated (r = 0.80), HSGPA was not included in the stratification process. ACT Composite and HSCentile were chosen because these criteria are used for admission to the university and provide an indicator of the academic preparedness of each student. For the purpose of this study and to simplify the stratification process, the HSCentiles were divided into quartiles (e.g., 1st quartile is 76-100), and the ACT Composite scores were divided into five separate score ranges. Quartiles were chosen as they are often used by the university to categorize students. The ACT categories were chosen for expediency. For example, thirty students with ACT Composite scores of 26 through 30 and HSCentiles in the 1st Quartile were randomly selected to be a part of each bootstrap sample. Eleven students within the same ACT Composite score range and the 2nd quartile were randomly selected for each bootstrap sample. Students without HSCentiles and/or ACT Composite scores were also included in the stratification process. For example, three students with ACT Composite scores of 31 through 36 and no HSCentile were randomly selected for each bootstrap sample. Table 2 presents the frequency of ACT Composite scores and HSCentiles of the 2012 LC cohort.

For the purposes of this study, three bootstrap samples were selected. The data from these three samples were averaged in the Bootstrap Samples column of Table 1. As this column indicates, the mean admissions characteristics of the combined three bootstrap samples very closely match those of the Fall 2012 LC cohort. This allows for better comparison of the 2012 LC and non-LC groups since they are roughly equivalent in academic preparedness.

It is worth noting that different variables (e.g., high school GPA), as well as different strata of HSCentiles (e.g., deciles) and ACT Composite scores (e.g., smaller or larger score ranges) could have been selected. Additional bootstrap samples could have also been selected for this study. Furthermore, more sophisticated statistical analyses could have been employed (e.g., a stepwise linear regression accounting for the variability associated with HSCentile and ACT Composite before evaluating the impact of the LCs). These changes may have yielded slightly different results. However, the purpose of this article is to provide a simple way for LC administrators to analyze benefits of LCs while accounting for self-selection bias. This is especially useful when LC administrators have limited time, resources, or training in data management and statistical methodology.

Table 2
Fall 2012 LC ACT Composite Scores and High School Percentile Ranks

	ACT Composite Scores						
High School Percentile Rank	11-15	16-20	21-25	26-30	31-36	No ACT Score	Total
1st Quartile	1	17	34	30	13	2	97
2nd Quartile		34	33	11	2	1	81
3rd Quartile	1	12	23	2			38
4th Quartile		1	4				5
No HSCentile	2	11	21	12	3	3	52
Total	**4**	**75**	**115**	**55**	**18**	**6**	**273**

Results

The mean university (i.e., non-transfer) cumulative GPA was computed for the LC and non-LC cohorts each fall and spring semester. Although summer coursework, if any, is included in the cumulative GPAs, summer cumulative GPAs were not evaluated in this study. Likewise, the percent of students who re-enrolled at the university each fall and spring semester was computed. These data are regularly provided to one of the authors to evaluate different programs at the university.

Figures 1 and 2 present the mean university GPAs and retention of the LC and non-LC cohorts for the Fall 2010 cohort, and Figures 3 and 4 present the same data for the 2011 cohort, respectively. As these figures indicate, the 2010 and 2011 LC students outperform the non-LC cohort for as long as six semesters (Figure 1) despite having very similar entry characteristics (Table 1).

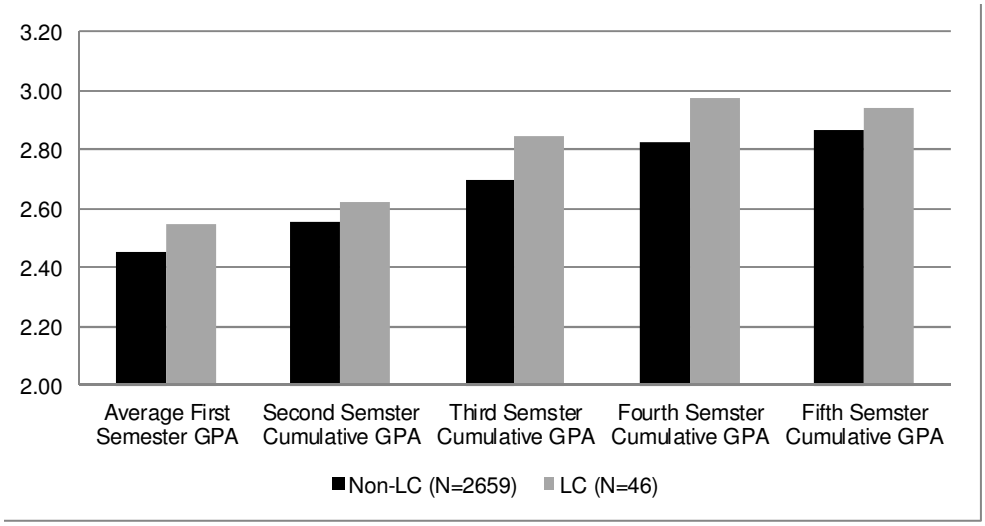

Figure 1. Mean University GPAs of the Fall 2010 LC and non-LC Cohorts

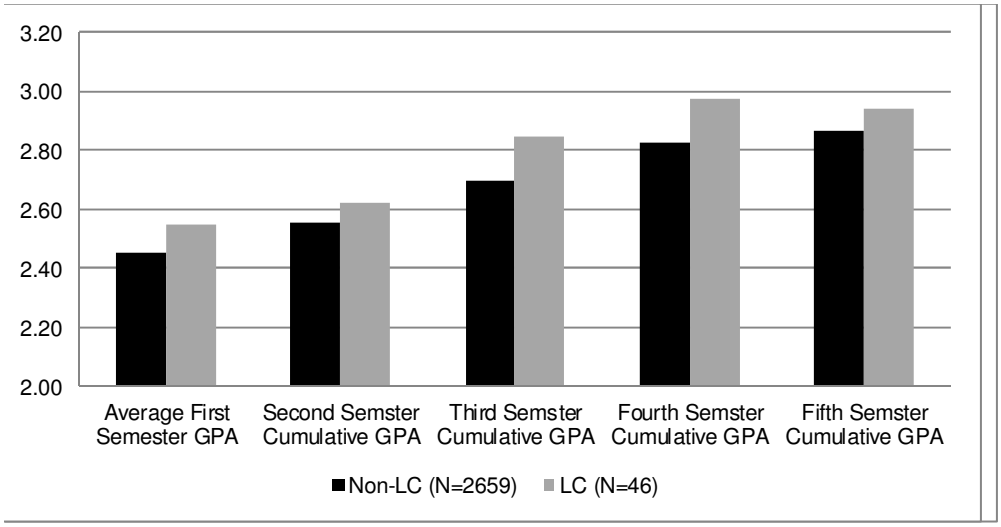

Figure 2. Mean Retention of the Fall 2010 LC and non-LC Cohorts

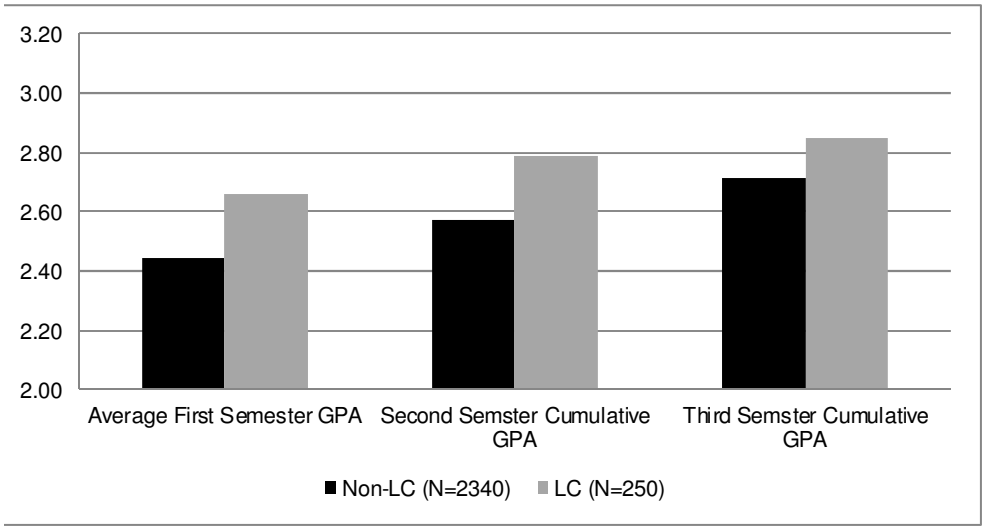

Figure 3. Mean University GPAs of the Fall 2011 LC and non-LC Cohorts

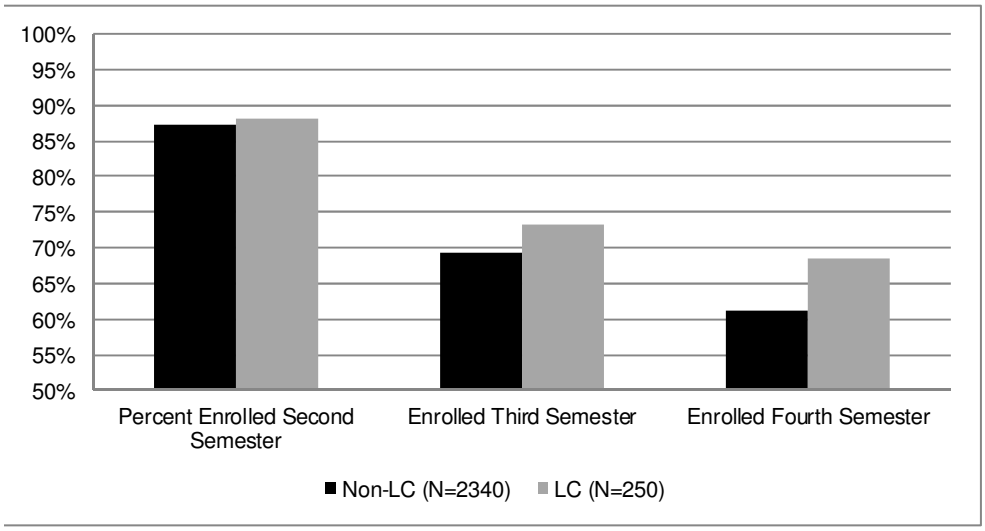

Figure 4. Mean Retention of the Fall 2011 LC and non-LC Cohorts

In Fall 2010 and 2011 the LC students were very similar academically to the non-LC students based on admission characteristics (Table 1). The Fall 2012 LC cohort was considerably more prepared academically than the non-LC cohort. Figures 5 and 6 present the mean GPA and retention rates of LC, non-LC, as well as the three bootstrap samples. Figures 5 and 6 indicate that the Fall 2012 LC cohort outperforms both the non-LC cohort and the bootstrap sample. This

effectively negates the argument that students who self-select into LCs outperform non-LC students based solely on their academic preparedness and strengthens the argument for the benefits of LCs.

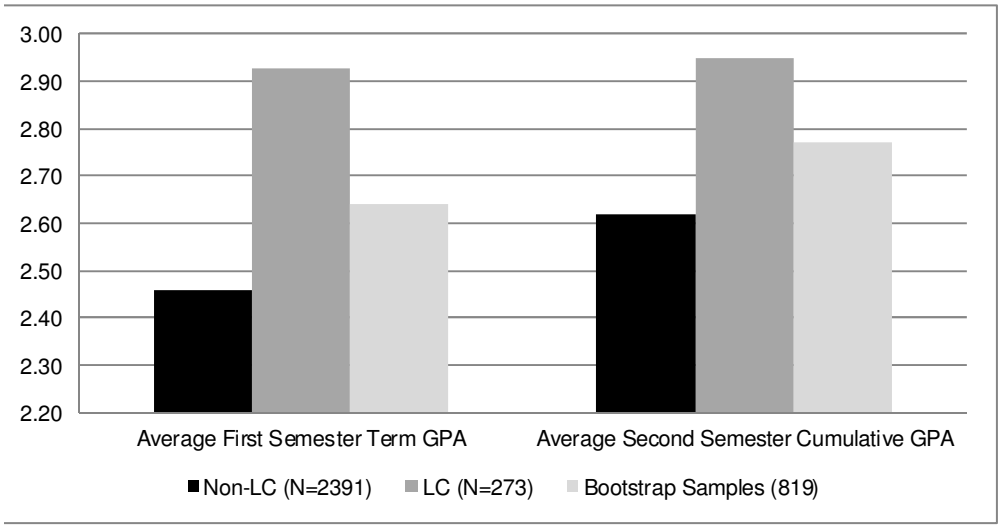

Figure 5. Mean University GPAs of the Fall 2012 LC, non-LC, and Bootstrap Samples

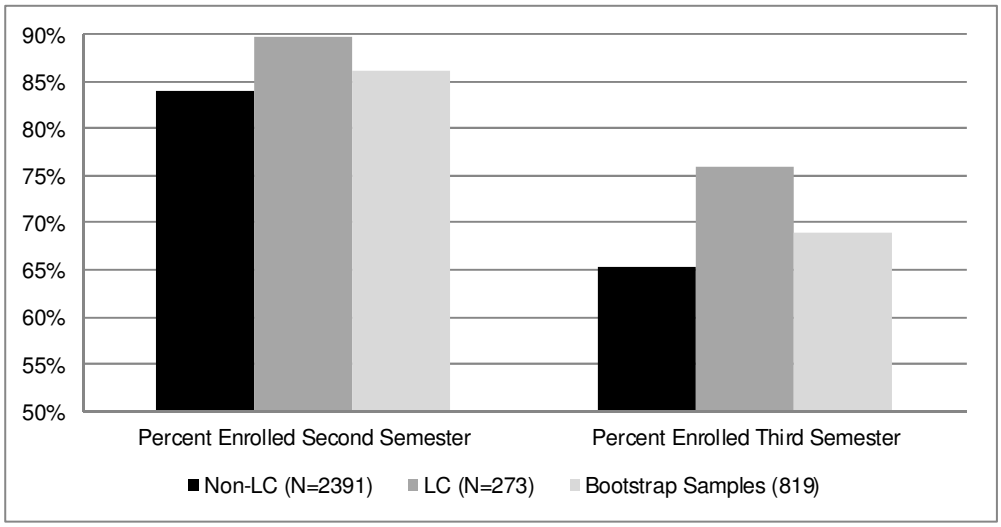

Figure 6. Mean Retention of the Fall 2012 LC, non-LC, and Bootstrap Samples

Limitations

There are several limitations to this study. The first limitation relates to the limited number of variables used to stratify the three random samples that comprised the bootstrap sample. To create the comparison groups, two variables, ACT score and high school percentile rank, were used. In doing so, quick and simple comparison groups were created that strongly reflect the academic level of the LC cohort. However, students also have a variety of other inputs that might affect their success upon entering college, including many that Zhao and Kuh (2004) control for in their multi-institutional study such as status (full- or part-time), age, gender, class, race/ethnicity, etc. Other factors non-academic indicators (e.g., finances, family issues, homesickness) may also influence the success or failure of students. Future research may choose to include additional variables for stratification.

A second limitation of the study is our inability to determine the exact intervention(s) of the LCs that result in increased student success. Students participating in LCs are subjected to multiple interventions, including the completion of MAP-Works surveys (early alert system), access to an upper class peer leader for academic and social support, and additional socialization opportunities such as the orientation event and other community building activities. Additionally, students in LCs often complete a first-year seminar course tied to the LC. Thus, it is possible that the complex combination of interventions results in increased student success (Pike, 2000; Pike, Schroeder, & Berry, 1997; Zhao & Kuh, 2004), a possibility that might be explored in future studies.

Furthermore, to gain a better understanding of what factors might contribute most to the higher GPA and retention of these students, one might consider employing various methods to gather additional information on academic success. Coupling the bootstrap samples with student focus groups or interviews might provide a more holistic understanding of the factors influenced academic achievement (Pike & Ouimet, 2009).

Conclusion

This study suggests that LCs positively impact student academic achievement and retention. By creating an equivalent comparison group through the bootstrap sample, this study indicates that any advantage realized by LCs students were not due entirely to self-selection bias. This methodology offers a simple and easy solution to address possible self-selection bias. Other campuses can employ similar methods to demonstrate the benefits of LCs to key constituency groups, including university administrators, academic advisors,

faculty, staff, students, and parents. Future studies could include additional variables (e.g., living on/off-campus, parental income, other campus involvement) to examine the impact of LCs in greater depth.

References

Astin, A. W. (1993). Student involvement: A developmental theory for higher education. *Journal of College Student Development*, *40*(5), 518-529.

Baker, S., & Pomerantz, N. (2000). Impact of learning communities on retention at a metropolitan university. *Journal of College Student Retention*, *2*, 115-126.

Brower, A. A., & Dettinger, K. M. (1998, November/December). What is a learning community?: Toward a comprehensive model. *About Campus*, *3*(5), 15-21.

Brower, A. M., & Inkelas, K. K. (2010). Living-learning programs: One high-impact educational practice we now know a lot about. *Liberal Education*, *96*(2), 36-43.

Cross, K. P. (1998, July/August). Why learning communities? Why now? *About Campus*, *3*(3), 4-11.

Gabelnick, F. G., MacGregor, J., Matthews, R. S., & Smith, B. L. (1990). *Learning Communities: Creating connections among students, faculty, and disciplines*. New Directions for Teaching and Learning, *41*. San Francisco, CA: Jossey-Bass.

Huerta, J. C., & Bray, J. J. (2013). How do learning communities affect first-year Latino students? *Learning Communities Research and Practice*, *1*(1), Article 5.

Inkelas, K. K., & Weisman, J. (2003). Different by design: An examination of outcomes associated with three types of living-learning programs. *Journal of College Student Development*, *44*, 335-368.

Inkelas, K. K., Vogt, K. W., Longerbeam, S. D., Owen, J., & Johnson, D. (2006). Measuring outcomes of living–learning programs: Examining college environments and student learning and development. *The Journal of General Education*, *55*(1), 40-76.

Kuh, G. D. (2008). High-impact educational practices: What they are, who has access to them, and why they matter. Washington, DC: Association of American Colleges and Universities.

Lenning, O. & Ebbers, L. (1999). The powerful potential of learning communities: Improving education for the future. ASHE-ERIC Higher Education Report *26*(6). Washington, DC: The George Washington University, Graduate School of Education and Human Development.

Matthews, R. S. (1994). Enriching teaching and learning through learning communities. In T. O'Banion (Ed.), *Teaching and learning in the community college* (pp. 179–200). Washington, DC: Community College Press.

Mooney, C. Z., & Duval, R. A. (1993) *Bootstrapping: A nonparametric approach to statistical inference*. Sage University Paper series on Quantitative Applications in the Social Sciences,7-095. Newbury Park, CA: SAGE Publications.

Pike, G. R. (1999). The effects of residential learning communities and traditional residential living arrangements on educational gains during the first year of college. *Journal of College Student Development, 40*(3), 269-283.

Pike, G. R. (2000). Assessment measures: Methodological issues in the assessment of learning communities. *Assessment Update: Progress, Trends, and Practices in Higher Education, 12*(2), 14-15.

Pike, G. R., Kuh, G. D,. & McCormick, A. C. (2011). An investigation of the contingent relationships between learning community participation and student engagement. *Research in Higher Education, 52*, 300-322.

Pike, G. R., & Ouimet, J. A. (2009). Assessment measures: Using theory and expert opinion, focus groups, and student interviews to improve campus surveys. *Assessment Update: Progress, Trends, and Practices in Higher Education, 21*(1), 7-9.

Pike, G. R., Schroeder, C. C., & Berry, T. R. (1997). Enhancing the educational impact of residence halls: The relationship between residential learning communities and first year college experiences and persistence. *Journal of College Student Development, 3*8, 609-621.

Shapiro, N. S., & Levine, J. (Eds.). (1999). *Creating learning communities: A practical guide to winning support, organizing for change, and implementing programs*. San Francisco, CA: Jossey-Bass.

Smith, B. L., MacGregor, J., Matthews, R. S., & Lindblad, J. (2004). *Learning Communities: Reforming Undergraduate Education*. San Francisco, CA: Jossey-Bass.

Stassen, M. L. A. (2003). Student outcomes: The impact of varying living-learning community models. *Research in Higher Education, 45*, 115-138.

Taylor, K., Moore, W., MacGregor, J., & Lindblad, J. (2004). *Learning community research and assessment: What we know now.* National Learning Communities Project Monograph Series. Olympia, WA: The Evergreen State College, Washington Center for Improving the Quality of Undergraduate Education.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago, IL: The University Press.

Tinto, V. (1998). Colleges as communities: Taking research on student persistence seriously. *Assessment Update 12*(2), 1-2, 12.

Tinto, V., & Goodsell, A. (1993). Freshman interest groups and the first year experience: Constructing student communities in a large university. *Journal of the Freshman Year Experience*, *6*(1), 7-28.

Tinto, V., & Love, A. G. (1995). *A longitudinal study of learning communities at LaGuardia Community College.* (Eric Document Reproduction Service No. ED 380178). Washington DC: Office of Education Research and Improvement.

Tinto, V., & Russo, P. (1994). Coordinated studies programs: Their effect on student involvement at a community college. *Community College Review, 22*(2), 16-25.

Zhao, C., & Kuh, G. (2004). Adding value: Learning communities and student engagement. *Research in Higher Education*, *45*, 115-138.