Mining Transactional Student-Level Data to Predict Community College Student Outcomes

Erez Lenchner

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY
2017

ABSTRACT

Mining Transactional Student-Level Data to Predict Community College Student Outcomes

Erez Lenchner

A longitudinal analysis of transactional data for an entire college cohort was mined from administrative student records systems to identify individual student behaviors and establish correlations between individual students' behaviors and academic outcomes. Conducted at one large urban community college, this study determined curricular peer association behavior between individual students, and also evaluated late registration and course schedule change behaviors. Findings demonstrated a strong correlation between these three behavioral patterns and a lasting influence on academic outcomes, such as: semestrial GPA and cumulative GPA, credit accumulation, persistence and graduation rates. Finding also indicated a correlation among the three behaviors themselves. Furthermore, conducting a longitudinal analysis of individual students made it possible to identify the temporal tipping-points which differentiated at-risk behavior from otherwise benign behavior. The intrinsic factors associated with individual students' behaviors were followed over a period of thirteen consecutive semesters. Mining Transactional Student-Level Data at the scale achieved in this study, when compared to traditional methods of data collection, provided the precision needed to determine the actual proximity among specific peers, and the identification of registration behavior patterns. The extraction of transactional data from the records of each student in an entire cohort resulted in a method of inquiry immune to the negative effects of student's non-response or selection bias. Complimenting previous research, this study provides a detailed descriptive analysis of those behaviors not only at the semestrial level, but also cumulatively across consecutive semesters.

This study demonstrates that curricular peer association can be measured directly from common, ubiquitous, transactional records. The rates of Peer Association among individual students was very dynamic: While the majority of students had some peer associations while enrolled, in the aggregate two thirds of students had no peer association (were soloists) at some point in time, while more than a quarter of all students were soloists for at least half of their entire enrollment period.

Soloists differed from students with peer associations. They were likely to be older, international students, African Americans, transfer students, or those entering fully prepared for college level coursework (no remedial coursework). Peer association was positively correlated, both in the semester in which it occurred and cumulatively, with: GPA, credits earned, and retention or graduation rates. These correlations to academic outcomes varied with the number of peer associations established, and the intensity of peer encounters.

The study revealed that nearly a quarter of all students practiced late registration at least once; and more than 10 percent have registered late multiple times during their studies. Nearly three quarters of students made modifications to their course schedule at least once after the semester began. Overall, two fifths of students changed their initial schedule every semester. These behaviors were unrecorded in previous studies that were limited in the evaluation of longitudinal behaviors, used subsets of students and were subject to non-response bias. Late registration and student schedule changes was correlated with lower semestrial and cumulative academic outcomes. Late registration behavior subsequently increased the likelihood of a student being a soloist. When compared to previous studies, the analysis conducted here not only accounted for academic, demographic and financial variables at baseline, but went on to perform updates at key points in time each semester to reflect changes over time. The exhaustive revisiting of the covariates each semester provided enhanced control to the ‘order of time’ influence. All covariates were re-measured each semester allowing to better evaluate the correlation of student behavioral indicators for a given semester, and cumulatively. This enhanced the study’s ability to account for common unobserved variables inherent to academic, demographic and financial attributes that might influence student outcomes correlated with peer association, late registration and schedule changes.

This study contributes to the literature by showing that peer association can be evaluated in the setting of an open admission commuter institution, and that peer association has consistent and positive correlation with academic outcomes. It provides new insights regarding the magnitude of late registration and schedule changes, as well as their negative immediate and longitudinal correlation with student outcomes. Further implications to community colleges’ faculty, administrators, researchers and policymakers, as well as future directions for research employing transactional level data are discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

My journey down the path towards becoming a researcher started when I met Elia Werczberger, Professor of Public Policy at Tel Aviv University. As a mentor Elia gave me the opportunity to work alongside him and learn the foundations of policy research, and the tools to navigate state, national and international statistical agencies. Next, Mark Scott at NYU inspired me through his outstanding teaching of advance statistical methods and his pioneering work in applied statistics. Marc always encouraged me use my creativity, dare to question established practices, and be persistent. I will always be grateful for his longstanding support and encouragement. I am also fortunate to have worked alongside Robert "Bob" Norman who further helped me address complicated statistical procedures and better communicate them to others.

For his tireless support over these many months I want to thank my Professor and dissertation chairperson, Prof. Thomas R. Bailey. He was very generous in providing steady encouragement, expert advice, and overall mentorship. Tom always pushed me question my own assumptions, consolidate my writing, and challenged me to expand my research methods, formulate broader conclusions and discern the implications which might inform the issues facing community colleges nationwide. I would like to thank my dissertation committee members: Clive Belfield, Kevin Dougherty, Douglas Ready, and David Weiman for their detailed input and advice in the dissertation review process.

I want to thank the Economics and Education faculty and staff, and the CCRC crew that made the doctoral path a great research and professional growth experience. In particular, I would like to thank Henry Levin, Tom Bailey, Kevin Dougherty, Regina Cortina, Judith Scott Clayton, Davis Jenkins, Jim Jacobs, Matt Zeidenberg, Olga Rodriguez and Peter Bergman for their insights, support, and advice throughout my research and professional growth. I am grateful to Gladys Perez-Mojica, Liz Carroll, and Sarah Prescott Phillips for their support and help navigating the TC system, schedules, deadlines and numerous bureaucratic obstacles. I have to thank many friends and fellows at the National Center for Education Statistics (NCES), Research Triangle Institute (RTI), and the Association for Institutional Research (AIR) for their longstanding support, and the endless times in which we brainstormed research ideas at different hours of the night and weekends, developed modules, evaluated policy research and teaching paths, and for their inspiration and ideas along the way.

A very special thanks to my many friends who provided ongoing help and moral support in order to persevere throughout my research. I want to show my appreciation for the input provided on numerous ideas and critiques of the manuscript as it evolved to: Henry Saltiel, Jeff Weintraub, Sandra Kinney, Kristina Powers, Kurt Gunnell and many more. For their long-term collaboration and support, I need to thank my fellows in the office of Institutional Research and Assessment, over the years, by (approximate) chronological order- Henry Saltiel, Jan Gau, Jeff Weintraub, Anatoliy Kadinskiy, Barbara Astone, Veron Sundar, Patricia Rivera, Nate Dickmeyer, Jenny Zhu, and Sam Michalowski. I am infinitely grateful to LaGuardia College's President, Dr. Gail O. Mellow, for her tireless devotion to supporting the mission of community colleges, and for sustaining a research environment that fosters a true culture of inquiry to help promote community college student success at the institutional and national level.

CHAPTER I

INTRODUCTION

In the study of factors that influence community college student outcomes, administrative records can play a vital role in uncovering inherent behavioral patterns. Using large data sets to conduct predictive analytics enables colleges to reveal patterns in students' actions and unearth relationships to some of the underlying associated student factors. Recent practices in academic analytics, such as using transactional student-level data and predictive analytics only dates back to approximately 2000 (Campbell, DeBlois, & Oblinger, 2007; Peña-Ayala, 2013, 2014). These practices have vastly expanded the ability of researchers to gather and analyze student-level administrative records. Transactional data analysis starts by gathering live and historical data stored in information systems supporting student learning or administrative transactional records (Baepler & Murdoch, 2010; Ferguson, 2012; Luan, 2002, 2007; Cristobal Romero & Ventura, 2013). Ready access to data enhances its value allowing for rapid turnaround, from data gathering and analysis to utilization, expanding the real time usage of such information (West, 2012). While this information is widely available, it is rarely compiled and used for the purpose of rigorous analysis, or to develop and assess services or outreach for at-risk students in community colleges.

Precisely the extent to which present day transactional data allows investigators to explore the effect of student behaviors on academic outcomes in community colleges has yet to be documented thoroughly. While studies have explored the relationship of individual behaviors to student outcomes, investigators have not focused on the temporal tipping-points which might help identify at-risk behavior from otherwise benign behavior (e.g. schedule changes that reflect on academic challenges, unlike ones that reflect on a student work or course block scheduling).

To that end this study begins to address this gap in data by examining the timing and frequency of various student behaviors. Previously, the behaviors associated with: late course registration; frequency of course changes; and departure from prescribed curricular pathways were correlated in many studies with negative student outcomes. On the other hand, behaviors associated with establishing social support networks; participation in new student orientation; and the maintenance of continuous enrollment or enrollment intensity in accordance with prescribed curricular pathways have been shown to contribute to positive student academic outcomes.

The primary motivation of this study is to identify evidence of student behaviors within the transactional data, (specifically--late registration, schedule changes and student peer association) and examine if correlations exist between those behaviors and academic student outcomes. The interest here in those specific behaviors stems from a desire to inform the lack of attention and detailed documentation in the community college literature. Understanding the effect of these behaviors when they occur, and cumulatively is crucial if they are associated with changes in student outcomes in degree programs:

Higher education researchers have visited late registration behavior before, but the extent to which community college students first exhibit late registration behavior at some point in time, or repetitively, has not been thoroughly explored. In previous studies it has been difficult for researchers to make comparisons because late registration has been ill-defined across institutions, and the sampling and the quality of data has varied widely. That being said, Community Colleges still operate with a lack information regarding the overall magnitude of late registration, or its effect on student outcomes. The present day ability to analyze transactional data allows investigators to document thoroughly the patterns of late registration and its association with student outcomes not only at the time of occurrence, but also cumulatively. It can provide faculty, researchers and administrators with better documentation of the phenomenon, its size and the potential implications of student lateness on their outcomes. Community Colleges would have an interest in understanding the magnitude of late registration behavior because it may affect decisions related to the allocation of teaching resources. Previous studies have suggested that permitting late registration carries negative consequences related to student academic outcomes. This creates conflict for colleges pressured to encourage accessibility to postsecondary education and also to meet their enrollment or financial targets. In recent years several university systems have either banned the practice or imposed limitations on late registration (Florida, 2012; O'Banion, 2012; SUNY, 2015).

Student initiated schedule changes has rarely been examined as an indicator of community college student outcomes. Although researchers, faculty and practitioners might have an interest in schedule change behavior, very little information is available regarding the proportion of students practicing schedule changes, nor its association with performance in academic programs. Even the direct extraction of data from administrative records can result in the misinterpretation of schedule changes because schedule changes are rarely reflected on student transcripts, and many changes are not even documented in college's registration transaction history - This study used a combination of sources and daily sampling to construct an audit trail of enrollment changes. Students' Transactional Data allows investigators to bridge the gap and document schedule changes that were previously not documented and not reported by students, or by the colleges. Schedule changes are of interest for student-level analysis, and for advisement and campus planning: At the student level, they may reflect on the students' organizational skills, ability to maintain a prescribed curricula, and external constraints to their academic progress that are not reported in other channels (e.g. work schedule, family commitments, etcetera). While it may reflect on students changing their interests once the semester begins, it may also indicate a low confidence level in a course or major materials. For academic advisors and faculty, a better understanding of schedule change behavior could provide invaluable information regarding this student behavior and allow to identify students who lack direction earlier in the academic cycle. It would also allow scholars to better evaluate the rate of schedule changes at community colleges, and create benchmarks for the evaluation of schedule changes for different risk groups (e.g. first generation college students, students who are in remedial coursework, etc.). At the planning level, schedule changes carry direct implications to the organizational settings of community colleges. For example, community colleges might need to offer more sections of key courses to enable students to adjust their schedule. As colleges cannot practice ‘overbooking’ of classes and classrooms it is crucial for researchers, faculty and practitioners to properly estimate the magnitude of student-initiated schedule changes. Describing the current patterns of schedule changes would allow practitioners to take them into account during the creation of an academic calendar, class scheduling and perhaps also in creating new processes for students to examine and select courses.

Since many studies have attributed student peer association with contributing to positive student outcomes in colleges (J. D. Angrist, 2014; Carrell, Fullerton, & West, 2009; Garlick, 2013; Lyle, 2007; Sacerdote, 2001, 2011; Zhao & Kuh, 2004; Zimmerman, 2003), the interest of this study is in the documentation of student peer association in the setting of commuter, community colleges, and its relationship to student academic outcomes. The majority of student peer association studies took place in highly selective four year institutions, and relied heavily on proxies (physical proximity, dormitory assignment, etc.) to determine that such associations existed. However, previous peer association studies rarely examined if curricular peer association indeed exists. If students did not in fact interact with the identified peers, the association to student outcomes may be simply overestimated. Mining transactional student-level data allows investigators to address this gap by documenting the direct student peer-association as reflected in the course registration records. It allows researchers to ask questions related to the efficacy of learning communities (e.g. do learning communities actually result in student peer association across other classes?), and it would fill gaps in the literature pertaining to the association of: the number of peer associations, and the intensity of peer associations students experience (and need in order to succeed) in their academic studies. Furthermore, this study is interested in the measurement of cumulative and point-in-time peer association in settings were random assignment is not available.

Statement of the problem:

The evolution of practices employed to identify community college's at-risk students may still be made more precise. Community colleges routinely classify at-risk students a single time, based solely on data derived at the time of admissions, and often only on a sub-set of students, those representing First Time Full Time (FTFT) admits. This static data fails to inform the fact that student's situation change over time. As new information systems have been widely adopted at community colleges, and new techniques for mining administrative transaction data have been developed over the last decade, those changes are traceable during the course of the student's enrollment. However, these data are rarely utilized by community colleges faculty and administrators to support students. As a matter of convenience, the attributes of first-time full-time students have been often carried over to all students, when a finer distinction of student attributes is likely to yield services better suited to at-risk students. There is mounting pressure on community colleges to improve student outcomes, and despite innumerable attempts to do so, relatively little change to retention, persistence and graduation outcomes has taken place. Given this pressure, exploring new methods to use student-level data during the course of enrollment will provide better insights for at-risk students. Focusing on actions traceable through transaction tables ensures that those insights would be applicable to virtually every community college that employs student information system, irrespective of a vendor and computing environment.

Research Questions and hypotheses:

This study addresses the following research questions:

(1) Does the use of student-level transaction data enhance the explanatory power of student-level behavioral indicators to predict community college student outcomes with regard to:

   a. Late Registration?
   b. Schedule changes?
   c. Peer effects?

$H1_a$: Student late registration contributes to reduced academic performance.

$H1_b$: Student-initiated class schedule changes made after the semester begins contributes to reduced academic performance.

$H1_c$: Evidence of peer-effect is associated with improved academic performance

(2) How are student-level behaviors related to the community college outcomes of:

   a. Semester to semester retention?
   b. Semester credits earned?
   c. Cumulative credits earned?
   d. Semester GPA?
   e. Cumulative GPA?
   f. Graduation?

$H2_a$: Student late registration and schedule changes will contribute to reduced retention rate; the evidence of peer-effect and participation in freshman seminar will contribute to increased retention rate

$H2_b$. Student late registration and schedule changes will contribute to lower levels of semester credits earned; the evidence of peer-effects and participation in freshman seminar will contribute to increased semester credits-earned

$H2_c$. Student late registration and schedule changes will contribute to lower levels of cumulative credits earned; the evidence of peer-effects and participation in freshman seminar will contribute to increased cumulative credits-earned

$H2_d$. Student late registration and schedule changes will contribute to lower levels of semestrial GPA; the evidence of peer-effects and participation in freshman seminar will contribute to increased semestrial GPA

$H2_e$. Student late registration and schedule changes will contribute to lower levels of cumulative GPA; the development of a network of peers and participation in freshman seminar will contribute to increased cumulative GPA

$H2_f$. Student late registration and schedule changes will contribute to lower graduation rates; the evidence of peer-effects and participation in a freshman seminar will contribute to increased student graduation rates.

CHAPTER II

LITERATURE REVIEW

This study is informed by research on student outcomes (retention, persistence and graduation) in community colleges. It draws on previous studies that employed various sources of data, including transactional data, to evaluate the relationship between student outcomes and student-level behavioral indicators. The review begins with an outline of the education data mining framework. Then, this review is divided into the following sections that are relevant to the current research. First, it examines several behavioral indicators associated with negatively effecting student progress - in particular, students' behavior related to delayed registration and schedule changes. Next, behavioral indicators associated with a positive effect to student progress are examined - in particular, student peer association and peer effects.

Each section summarizes the common definition of the problems or behaviors in question, the proportion of students presenting the behavior or symptoms in question, the common effect measured on student outcomes and the common policies used in addressing such behaviors. It concludes by mapping the main research methods and findings in each area, and examines the limitations of previous research in the assessment of behavioral student outcomes.

Student-Level Indicators in Administrative and Transactional Data Studies.

To effectively study student-level behavior in post-secondary education requires researchers to evaluate human behaviors by compiling transactional data in the usage of modern Information Technology systems in transactional data analysis has been a growing trend since 2000. Despite the availability of institutional student data systems, very few studies have employed detailed transactional data to analyze student behaviors and address challenges pertaining directly to the student experience in college. Transactional data has been primarily used to categorize students (pre enrollment or during their first semester), and to perform descriptive analysis useful in curriculum planning studies.

In the field of community colleges research, Bahr (2010) employed student-level transactional data and clustering techniques to create a typology of community college-going students. Using retrospective data from the California Community College registration system, Bahr found that students fell into six clusters: transfer, vocational, drop-in, noncredit, experimental, and exploratory. The clustering did allow the classification of students into descriptive groups at entry but it did not reclassify students during their enrollment period in the community college system. As such, it did not account for changes in student growth or attitudes despite the availability of ongoing updates to descriptive student data available during the study period. Various aspects of student change were not taken into account, specifically, Bahr notes that the measurement addresses student's initial enrollment and focused on first time students. Furthermore, Bahr notes that his study does not account for student remedial preparation level at entrance, nor does it account for students' completion or exiting of remedial coursework. In addition, Bahr notes that his study, as well as other studies classifying students into groups/clusters (e.g. Adelman (2005); Hu and McCormick (2012)) suffer inaccuracies revealed by the variation in findings derived through survey instruments administered to measure student aspirations , and those of actual behaviors reflected by student records. Even widely used surveys, such as NSSE (National Survey of Student Engagement) and CCSSE (Community College Survey of Student Engagement) which are held to be reliable indicators for student attachment to college need to take into account the potential bias in student self-reporting. Using existing transactional datasets allows for enhanced analysis of student-level outcomes, by reclassifying students in real time, and addressing their actual behaviors in lieu of their stated goals.

There are examples of studies that have employed techniques to update student-level records during their enrollment using transactional data. Knauf, Sakurai, Takada, and Tsuruta (2010) employed transactional registration data to examine student curriculum planning (specifically, in the engineering professions) and to evaluate a pathway that optimizes student learning and success. Through the development of a "Dynamic Storyboarding System", they evaluated student's learning processes and evaluate majors' curricula. This study was mainly descriptive in nature, and did not evaluate the quality of each student pathway nor its potential contribution to student success. In a more recent study of theirs (2013), Tsuruta et al. expanded on their earlier study (2010) which had focused exclusively on curriculum planning in order to answer two broader questions:

(a) What do the successful students' paths [in college studies] have in common?

(b) What paths distinguish more successful students from those of less successful students [in their college studies]?

While Tsuruta et al. (2013) show that successful students may take certain pathways that differ from other students, their studies did not take into account out-of-classroom student behavior. They did not examine student's development of peer networks, registration patterns or relationship between the required orientations and seminars for freshman and actual attendance. Furthermore, their analysis focused exclusively on students who are highly prepared academically (admitted to competitive engineering programs), and did not examine the settings of open-admission institutions. Third, given the focus on a subset of programs in engineering they could not account for program-level effects, which may exist alongside with the selected pathway. Fourth, due to the nature of the tested programs, the number of iterations and pathways a successful student may take is limited. Therefore, their findings may simply reflect on students' ability to follow a well prescribed sequence in lieu of making individual decisions. Meanwhile, Tsuruta et al. (2013) suggest that employing transactional data could be fruitful in terms of computing the mined results. Specifically, they note that the adaptation takes place in relation to both the educational history of the considered students (so that the longer the student is part of a study the more information is gained about her performances), and the database, whose data is dynamically updated by the students' study results reflected on the transactional data (at the cohort level), also increases the accuracy of data mining over time. While previous studies (Bahr, 2010; Hu & McCormick, 2012) used profiling data with little or no updates, Tsuruta shows that a profiling approach that is dynamic (and takes into account recent history of the student using transactional data) is likely to improve at risk models' predictive accuracy over time.

Employing those practices in the settings of an open-admission institution will greatly benefit advisors, administrators and faculty struggling with limited resources to address the needs of at-risk students.

Drawbacks of Previous Studies: The Need to Focus on Transactional/Time-Based Data

Studies relying on transactional data analysis borrowed techniques, methods and tools from DM, artificial intelligence and other closely related fields to have a particular focus on theoretical and practical education data challenges. While current studies have drawn from administrative datasets, the current study adds to the literature by using transactional data. Transactional data better presents student behaviors and interactions with the college system(s), and may better reflect on student-level affects than direct response to surveys. It accounts for all student-level transactions (e.g. adding, dropping, replacing courses) throughout their enrollment history at a college. Student-level behaviors (derived from the transactional data) can serve as indicators for their progress in the degree path, or as potential indicators of departures from the degree path. Specifically, the ability to monitor those indicators on real-time (semestrial) basis, allows to improve community college's administrators, advisors and faculty to focus limited resources on particular groups of at-risk students.

Late Registration

A review of the literature associated with late registration shows numerous definitions were used to identify the behavior. Researchers have differentiated late registration from other forms of registration in terms of a specific point in time, or a range of dates, relative to the first day of classes. Of these, the majority describe late registration as any registration occurring on the first day of class or anytime afterwards (Belcher & Patterson, 1990; Hale & Bray, 2011; Hiller, 2005; Mendiola-Perez, 2004; O'Banion, 2012; Safer, 2009; Schmidt, 2004; Sinclair Community College, 2005; Sova, 1986; Zottos, 2005). In Moore, Shulock, Ceja, & Lang (2007) only those registering for over 20% of their classes, on or after the first day of classes, were sampled. Other studies varied as to the range of dates sampled: during the first three days of classes (Neighbors, 1996); during the first week of classes (Diekhoff, 1992; Perkins, 2002); first eight days of class (Street, 2000); first twelve days of classes (Chilton, 1964; Parks, 1974); first day of class until the last day for adding courses (Diablo Valley College, 2004); three days before through eight days after the start of the term (Stein, 1984); one week after classes began and later (Angelo, 1990); two weeks before through twenty-one days after classes began (Cornille, 2009). In two other studies (Goodman, 2010; Peterson, 1986) limited the sample to only those students who registered for all their classes on or after the first day of classes.

Beginning with the earliest study found on late registration (Chilton, 1964), and overwhelmingly in those that followed, the literature suggests that late registration is negatively associated with student academic outcomes. There are few underlying reasons for the negative effect: Overall the findings indicate a late registrant student is more likely to withdraw or fail than students who register on time (Roueche & Roueche, 1993; Sinclair Community College, 2005; Sova, 1986; Tincher-Ladner, 2006). Most obviously the negative influence stems from the student's need to complete the course requirements for an entire semester in less time than required. They might begin a course without having completed an early assignment, missed a course orientation explaining the syllabus or was absent during assignment to working groups. Consequently the late registrant is set in a position where he or she needs to acquire the foundation knowledge for a given course retrospectively. Frequently, limited course seating compromises the late registrant into taking a class with a non-preferred professor and/or at an inconvenient time. Although barely addressed in the literature it would be fair to assume that arriving late to class connotes a negative social stigma perceived by both the professor and classmates. Lastly, college support resources are not geared to assist with the issues facing late registrants (Summers, 2000a).

Late Registration Studies

Despite the prevalence of late registration practices permitted at the majority of open-enrollment colleges, few rigorous studies have been conducted to determine the effect of late registration on postsecondary student outcomes. Method, population, study-period, and the definition of characteristics defining a late registrant varied widely.

Studies varied greatly in size by the number of students, semesters and institutions sampled. The largest late registration study found, (Moore et al., 2007), followed 260,215 degree-seeking students at 109 California Community Colleges from 1999-2006. The majority of studies have been conducted on relatively small samples, between 100 and 1,700 students, tested over a single semester, and confined to a single campus, or single academic major. Examples include: Diekhoff (1992) with an average N=110 students per year, Ford, Stahl, Walker, and Ford (2008) with N=253, Smith, Street, and Olivarez (2002) with N=251, Sova (1986) with 1,673, and Summers (2000a) with N=1,365.

Typically, studies sampled students for only one semester at a single institution by either using a randomly selected sample (Angelo, 1990; Chilton, 1964; Diekhoff, 1992; Keck, 2007; Mannan & Preusz, 1976; Neighbors, 1996; Parks, 1974; Peterson, 1986; Street, 2000) or the entire student body (Belcher & Patterson, 1990; Goodman, 2010; Sinclair Community College, 2005; Stein, 1984). Others sampled all students at one college over longer periods of time: Cornille (2009), three fall semesters; Mendiola-Perez (2004), three years; Perkins (2002), two fall semesters; and Safer (2009), two years. Hale (2011) studied all students at three rural community colleges over four years, and Zottos (2005) sampled from nine campuses in one large urban community college for one spring semester. Several studies limited their sample to subpopulations: McWaine (2012), studied only African American males at one college over three years; Diekhoff (1992), studied introductory psychology students he had instructed in 50 individual classes over 14 years; Schmidt (2004) sampled only financial aid recipients for one semester, and Sova (1986) which sampled in precollege English and introductory English classes at one community college.

Findings of Late Registration Studies

The varied methodologies used for studying late registration has not surprisingly generated a varied set of results. All studies indicate that late registration is prevalent when the practice is permitted. Studies varied in the proportion of late registrants. Typically, the outcomes examined were limited to course completion rate, GPA, and persistence into a subsequent semester, with limited control for additional covariates.

Among late registrants, students were more likely to be males, African Americans, and Hispanics (Keck, 2007; Summers, 2000a), non-traditional students, and students not receiving financial aid (Summers, 2000a). Summers (2000b), found these outcome variables were associated with late registration even after controlling for the student characteristics of age, gender, ethnicity and academic intent.

In the two studies with the largest sample-size, Moore et al (2007) with 109 campuses, and Zottos (2005) with 9 campuses, reported one out of four courses (24% and 27% respectively) were added through late registration. In a broad sampling "11% of CCSSE [Community College Survey of Student Engagement] respondents (26,828 of 238,504) and 8% of SENSE [Survey of Entering Student Engagement] respondents (2,629 of 34,266) say they registered after the first class session for at least one class" (Center for Community College Student Engagement, 2012, p. 13).

Generally, late registration is predictive of earning lower GPA, and increased course withdrawals (Freer-Weiss, 2004; Hiller, 2005; Moore et al., 2007; Roueche & Roueche, 1993; A. B. Smith et al., 2002; Summers, 2003); and students were less likely to persist in school in subsequent semesters (Freer-Weiss, 2004; Johnston, 2006; A. B. Smith et al., 2002; Summers, 2000a).

Limitations of Previous Studies of Late Registration, and the Deficiencies that will be Addressed Utilizing Student-Level Transactional Data

The studies of student late-registration to date suffered from the following limitations that can be overcome using time stamped student-level transactional data:

First- lack of a unified, replicable definition of late registration. Among the studies reviewed comparison was made difficult owing to the lack of a standard definition of late registration. The lack of a standardized definition was, in part, due to the data collection methods (surveys vs. administrative records); lack of synchronization with the registration cycle of the institution; and usage of data in special or specific programs that may not be applicable to an entire college. The result was a limited ability to generalize conclusions regarding late registration. The lack of standardization in data sources documenting late registration, and in the data collection procedures, resulted in an inability to test and/or replicate the results of previous studies.

Second, limited investigation of the longitudinal impact of late registration. Most late registration studies did not look into the longer term affects (beyond one semester or academic year) on students' academic outcomes in terms of semester to semester persistence, early departure or graduation. In short, previous studies fell short on the evaluation of students' longitudinal outcomes, and as such could not be used to develop predictive models.

Third, lack of control for detailed covariates and dynamic changes over time. Previous studies' reliance on a limited set of covariates produced results with weak explanatory power compounding the effect of short study period. Specifically, none of the studies reviewed updated baseline data (such as: age, remedial status, financial attributes, academic attributes (major, course load, equated credits)), or behavioral indicators to account for change in a student's profile over time. As those covariates change and update regularly, it carries implications to the student profile and composite risk (e.g. financial aid dependency status often vary). It is possible that delayed registration may reflect on student's external constraints. Taking into account those constraints using covariates would better inform decision makers regarding the impact of the late-registration indicator compared with other constraints, such as availability of funds for tuition.

Fourth, the studies did not use late-registration behavior as a risk indicator for future performances nor as a possible 'proxy' for students' unobserved qualities. This can be attributed to the inability to correlate inter-semester patterns, financial and academic covariates, and other additional behavioral indicators such as: change of major, class attendance, stop-outs, et cetera.

Mining student-level transactional data in the evaluation of student late registration allows to overcome those challenges:

First, transactional (time stamped) data enables the researcher to map the full registration cycle of students, and evaluate the cycle against the transactional data at the individual (student) level. It then takes into account the established key-dates structure common for all colleges (e.g. first day of registration, first day of classes, and last day of a semester). Through the comparison of this information, transactional data enables a clear mapping of each student's late registration occurrences and patterns (e.g. a single course or multiple courses, or single instance or chronic behavior; early or late registration). As a result, employing the transactional data allows standardization of a late registration definition that can be verified, replicated and generalized across institutions.

Second- the student-level transactional data includes in the evaluation of late registration using virtually all students in the target group, and not just a subset of students who filled out a survey or have a detailed record available. It enables the retrieval of all registration records from the transaction files, including temporal information related to each action, which ensures full coverage of the student group.

Third, because the student-level transactional data contains all registration activities and provides a time stamp as to their occurrence, it the impact of late registration indicators can be evaluated of beyond the span of the semester or the academic year in which it took place. It can do this by linking the student record across semesters, and not through direct data collection (which is subject to the limitations of data completion and accuracy through surveying). Transactional data enhances the researcher's ability to document the behavioral indicator across time periods that is limited through other methods. Furthermore, transactional data allows the researcher to establish different weights to recent behavioral indicators compared with historical behaviors of the subject (e.g. schedule changes in a recent semester may provide a better indication than changes that took place two years ago).

Fourth- through student-level transactional data analysis the researcher has enhanced abilities to control for covariates and changes in a student's profile, and gains control for changing covariates over time, allowing for improved explanatory power within each semester during the longitudinal study period. Employing transactional data contributes directly to the models' ability to control for student-level qualities that impacts both positive and negative educational outcomes (Knauf et al. (2010)). It enables the researcher to account for a detailed academic and financial profile of the student which was not available for testing in previous studies. As such, studies employing student-level transactional level data can provide better proxies for students' unobserved qualities which may be reflected through a late-registration indication.

Schedule Changes (Course Shopping)

After a student registers for classes, either as an early or late registrant, they often make changes to their schedule by adding and or dropping courses. Rules governing these schedule changes vary by institution, but most schools allow these changes to occur before classes begin, and then for a short period (first one or two weeks) into the semester. Not all schedule changes are student initiated: Between 10% and 35% of schedule changes are due to course cancellations, time changes, departments adding courses, or mistakes made during the enrollment process or in advising ((Moran, Bausili, & Kramer, 1995; Morris, 1986) as cited in (Hagedorn, Maxwell, Cypers, Moon, & Lester, 2007)). With the exception of Broadbent (1975) and Hagedorn et al. (2007), the examination of course schedule changes have not been specifically targeted within student behavioral studies. Although there have been many course registration behavior studies, few rigorous longitudinal studies have been conducted to holistically examine the relationship between: individual student characteristics; course registration behavior (early -vs- late registration, late registration due to late admit, schedule changes (drop/add), and change of major); and various academic outcomes (i.e. GPA, credits attempted, retention, early transfer, etc.).

Student-initiated schedule changes are likely to be negatively associated with student academic outcomes as they occurred in response to dissatisfaction with professor or course content, level of difficulty, and inability to keep up with class assignments (Broadbent, 1975). More recently (e.g. Conklin (1997); Diablo Valley College (2004)), studies have reported that student-initiated schedule changes are likely to reflect on work schedule conflicts, time/inconvenient course schedule, external family or personal problems, concerns with academic grading and/or the instructor. These schedule changes indicate a need to accommodate external constraints to study during the semester (in lieu of planning ahead), falling behind with assignments, and adjusting to the expectations and requirements of new instructors. Incidentally, schedule changes may result in financial pressure through additional fees and/or tuition charges: Colleges typically employ a partial tuition refund policy for students who conduct schedule changes mid semester. Hence, a student who conducts schedule changes regularly may effectively pay a higher tuition rate, which will not be covered through student aid (if applicable). Premeditated schedule change among those students who intentionally perform 'course shopping' reflects an attempt to avoid courses with challenging academic requirements or the inability to predict external constraints such as work and/or family schedule, indicative of a lower commitment to the academic curriculum (Hagedorn et al., 2007). Conversely, academically successful students have fewer schedule changes, fewer course drops, and when they did conduct schedule changes they conduct them very early (within the first week of the add/drop period) (Summers, 2000).

Schedule Changes Studies

Three frequently cited studies that surveyed community college students' registration behavior found that the primary reasons given for course shopping (drops and adds) were the availability of courses offered or personal matters external to college. In addition, reasons such as: dissatisfaction with course content, level of difficulty, and inability to keep up with class assignments, where also cited. (Broadbent, 1975; Grimes & Antworth, 1996; Thomas-Spiegel, 1997).

Hagedorn et al. (2007) identified two main kinds of course shopping patterns and referred to them as: 1.) cyclic shopping (the pattern of dropping a course and adding another in its place) and, 2.) bulk shopping (signing up for more courses than the student expects to complete with the expectation of dropping some later in the semester). Fifteen percent of students practiced regular bulk and/or cyclic shopping (2007, Tables 3 and 7). Within these two behaviors cyclic shoppers were more likely to have lower GPA, dropped courses early in the semester, and then were more likely to withdraw from or fail their other courses. Approximately one quarter of all students practiced some form of occasional drop/add behavior which was not associated with lower academic performance.

To investigate why students changed schedules, and which courses were changed most often, Conklin (1997) administered a survey to all students requesting to drop a course at a large suburban community college in a five year study (spring 1988- spring 1993). The average rate of courses dropped during the study was consistently between 15%-16% of total enrollments. The five most frequent reasons respondents gave for dropping a course were: Work schedule conflicts (33%), Bad time/inconvenient (21%), Personal problems (18%), Too hard/bad grades (16%), and Disliked instructor (12%).

Several studies though have been conducted at community colleges which produced findings related to schedule changes in the context of a broader study of student registration behavior. In two such studies which were heavily cited in the literature a relationship between schedule change behavior and academic outcomes of semester-to-semester persistence and academic performance (GPA or credits earned) were found to exist (Diablo Valley College, 2004; Summers, 2000a).

In a descriptive study of student registration behavior on academic performance conducted at Diablo Community College (2004) all 122,000 students who enrolled from fall 2001 through spring 2003 (3 years) were analyzed. This study defined schedule changes as early-late registration and often referred to the practice as “course shopping”. The most cross-cutting finding was that students who were changing their schedules were outperforming other students. On average, 19% of the student population made changes to their class schedules. The Highest rate of course change, 25%, was among transfer-bound students (primarily full-time, young, White, Asian, or international students), while the lowest were students described as having a personal goal of lifelong learning at 8%. Overall, full-time and young students made the most schedule changes at 89% and 29.6%, respectively. Among the general student population the highest course success rates (73.7%) and course retention rates (83.4%) belonged to students whose registration strategy was to register as early as possible, and then make schedule changes during the late registration period (Morris, 1986); the lowest were late registrants at 63.5% and 77.1%, respectively. The same pattern of higher success rate was seen among remedial students, with schedule-changers having higher course success (66.7%) and course retention rates (82.2%) when compared to late registrants at 57.3% and 76.5%, respectively.

Summers (2000) conducted a rigorous examination of enrollment behavior for all first-time/full-time students who first matriculated at a rural Midwest community college. Individual cohorts were followed for their first fall and spring semester during three years (Fall 1994, Fall 1995 and Fall 1996). Summers collectively measured the number of schedule changes as any drop, add or section change; and differentiated an early schedule change as occurring any day prior to the start of the fall semester through the first seven days of the fall, and all others as late schedule change. Summers came to the following three conclusions: First, students who persisted from their first fall semester into their first spring semester averaged statistically significant fewer changes to their first fall semester course schedule (2.47 course drops) as compared to students not enrolled for the spring semester (4.18 course drops); secondly, students enrolled in the spring semester averaged significantly fewer course drops (1.32) than students not enrolled for the spring semester (3.03); and lastly, students enrolled in the spring semester averaged making a larger proportion (54.4%) of their schedule changes early (within the first week) of the add-drop period. White students, female students, and students who were not eligible for financial aid were significantly more likely to change their schedule during the early add-drop period than other students.

Summers (2000) found a correlation between schedule change and GPA, and GPA was in turn found to be the most significant factor associated with attrition. This strong relationship between GPA and attrition is consistent with Bean and Metzner's (1985) model of non-traditional student attrition, and also the structural equations model of total effects on persistence by Cabrera, et. al. (1993). When holding all student characteristics constant, and controlling for: number of course drops, when a student initially enrolled for the fall semester, and number of course adds; Summers found that the odds of persisting from fall to spring semester dropped 50.2% for each additional course dropped during the fall semester, and increased by 52.8% for each additional course added during the fall semester. In his best model to predict student fall to spring semester persistence, which controlled for: number of course drops, number of course adds, date student initially enrolled; and fall semester GPA and fall semester course completion, Summers determined for each additional course dropped by the student in the fall semester, the odds of enrolling for the spring semester decreased by 24.5%.

Limitations of Previous Studies of Schedule Changes, and the Deficiencies that will be Addressed Utilizing Student-Level Transactional Data

Previous studies of schedule changes faced the following drawbacks:

First, most registration behavior studies uncovered various associations, such as late registration, within the context of broad registration behavior studies, but did not treat schedule changes as a central focus. Given the constraints on data collection, previous studies faced challenges in the segmentation of schedule changes information from normal registration activities, and did not have access to document a student's full registration path. Compounding the paucity of results was the ill-defined and confusing array of ways in which late registration was defined. As a result, studies were limited in their ability to create and align a single definition of schedule changes that may be applied across colleges.

Second, the treatment of schedule changes in studies varied widely in terms of study period, and generally were conducted over very limited timeframes making it either impossible or ill-suited to examine the longer term effects (beyond one semester or academic year) on students' academic outcomes, such as long term retention and graduation trends. As such, previous studies fell short on the evaluation of students' longitudinal outcomes and therefore could not be used to develop predictive models.

Third, previous studies of schedule changes used information collected directly from students or administrators, and were not able to corroborate the records accuracy. In addition, direct data collection was subject to the constant decline in participants' response rate and response quality.

Fourth, previous studies fell short on their ability to accurately differentiate student-initiated schedule changes versus system or campus schedule changes due to limited ability to examine the change record. For that reason, they were subject to misclassification of the change cause, and likely caused an under-reporting of schedule changes.

Fifth, previous studies rarely used schedule changes or 'course shopping' as a risk indicator for future performance, nor was it possible to establish a 'proxy' for students unobserved qualities. In particular, it is possible that previous studies using surveying techniques were not able to link the reported behavior to a cause and effect relationship to the students' academic record.

Employing student-level transactional data allows researchers to overcome those deficiencies, and provided important advantages to the current study:

First- the student-level transactional data enables the researcher to map the full registration cycle, and evaluate it against the transactional data for each student. Through the standardization of key points in time , which could be applied to any college (e.g. first day of registration, first day of classes, and last day of a semester), it enables the mapping of schedule changes. Analysis that combines transactional data with 'key points in time' enables colleges to create a single definition of schedule changes that can be compared across colleges and provide a clear overview of schedule changes that was not previously available. It removes constraints of previous studies that relied on partial mapping of a registration cycle (and were sensitive, for example, to the definition of schedule change periods) and employs a consistent definition of schedule changes for all students in the sample.

Second- student-level transactional data allows for the evaluation of schedule change indicators beyond the course of a semester or an academic year, which requires a linkage to the student record (and not direct data collection). It enhances the researcher's ability to document behaviors across time periods, which is limited through other methods. Consequently, it allows the researcher to provide different weights to recent behavioral indicators compared with historical behaviors of the subject (e.g. schedule changes in a recent semester may provide a better indication than changes that took place two years ago).

Third, when evaluating student-level transactional records against the key dates, one can provide an accurate (virtually complete) mapping of a student schedule changes. It can be achieved because transaction information derived from the transactional data contains both individual (student) record 'stamp' and time record 'stamps'. Since the transactional information is associated with numerous additional systems (in particular, student billing and transcripts) its accuracy is high and updated constantly. Such a level of accuracy and redundancy of record checking is superior to the ones employed in existing schedule changes studies.

Fourth- mining student-level transactional data, allows to examine not only if a schedule change place took place but also the initiating mechanism (e.g. student, scheduling, bursar, etc.). It increases the accuracy of the information collected, and overcomes the challenges noted in previous studies regarding the completeness and accuracy of schedule changes.

Lastly, the richness of student-level transactional data greatly enhances the ability to recognize patterns of course schedules; in particular it allows to determine if a course schedule changes is a regular pattern for the student and provide direct evidence for 'course shopping' or 'bulk shopping', and estimate the magnitude of those behaviors on student outcomes.

Peer Effects

Although peer effects have been studied in both secondary and post-secondary education with relation to various academic outcomes, the main body of the literature on peer effects in post-secondary education is focused on traditional age, four-year, residential college students. More recently, studies have expanded the research to encompass commuter students, open admissions institutions and community colleges. In community colleges and commuter institutions this research has been centered on peer effects related to learning communities.

Student peer association is correlated with student outcomes. Indeed, most people have a powerful intuition is that peers matter in connection with a positive association between students and academic outcomes (J. D. Angrist, 2014; Griffith & Rask, 2014). Related to the potential benefits of peer association there has been a call to promote social relationships among non-traditional students in order to help them feel more comfortable in college and provide them with access to information that can ease their path toward a degree (Karp, 2011). Peer associations may influence student outcomes for several reasons. First, several peer effect studies noted that proximity to peers influences student (and individual) outcomes by gaining knowledge from their experience in similar settings. For example, students who have peers who went through a similar course or had similar experiences may benefit from their knowledge and understanding of the course (Zimmerman, 2003). Similarly, due to the inherent equality and social context of being a peer, students may be better positioned to interact informally to transfer knowledge (Winston & Zimmerman, 2004). Peer association may result in relationships which support the sharing of course materials, knowledge transfer, and support for events such as absenteeism. Having peers helps students get to know each other better and more quickly, which in turn leads to the development of social and academic support networks crucial to their progress in the degree (Visher, Schneider, Wathington, & Collado, 2010). When students form peer associations, they are more likely to get involved in collaborative learning (Vincent Tinto, 1993, 1997; Visher, Wathington, Richburg-Hayes, & Schneider, 2008). Collaborative learning, study groups and mutual support will enhances student interaction, promotes a sense of belonging, encourages time spent on studying academic coursework and academic skills (Engstrom & Tinto, 2008;

Rocconi, 2011; Visher et al., 2008). Ties to the college and academic program are also strengthened by peer relationships which results in lowering the probability of early departure. It further provides the student with opportunities to gain access both to informal in school knowledge (e.g. internships, scholarships, course openings etc.) as well as network for job searches, social benefits, and general support.

Newcomb's Model of Peer Influence

Initial research on peer effects in postsecondary education stems from the work of Newcomb (1962) and Newcomb, Brown, Kulik, Reimer, and Revelle (1970). In these studies students were seen as members of groups, and that those groups exerted social and cohesive influence on its members, such that individuals develop expectations and reliance of each other, as well as favorable attitudes towards one another. From this perspective, peer influence can be attributed to changes in students' attitudes rather than from the intrinsic attributes or skills that individual members possess (Newcomb, 1962; Newcomb et al., 1970; Weidman, 1989). Newcomb (1962) believed that this attitudinal change took place through the sharing of group norms; and was contingent upon the level in which the groups' messages took place in isolation of the external environment, within moderate-sized groups, and among members who shared common backgrounds. The amount of time students share in close proximity strongly influences these peer effects (Newcomb, 1962; Pascarella & Terenzini, 2005). For that reason the conditions for peer influence should be most prevalent within residential colleges, yet Newcomb (1962) and Berger and Milem (2000) point out that colleges can foster environmental conditions that enhance peer group influence, so that such affect can be utilized in a positive manner.

Previous Studies of Peer Effects

The general consensus within the post-secondary literature is that peer effect is associated with small yet positive effects on academic outcomes such as GPA at traditional residential institutions (Foster, 2006; Sacerdote, 2001, 2011; Zimmerman, 2003). The effect may differ by academic preparation level and previous achievements; and the effect of 'social' peers seems stronger than the effect of 'randomized' peers (Foster 2006). There was also little effect on decisions regarding academic major or career-choice patterns (Sacerdote, 2001). To a greater extent, peer-effect was found to have significant influence on personal behavioral decisions, in particular extra-curricular activities, and also on the level of academic effort outside of the classroom (Sacerdote, 2001; Zimmerman, 2003).

There are some exceptions. Peer effect studies conducted at the U.S. Army (Lyle, 2007) and U.S. Air Force academies (Carrell et al., 2009), where intense peer-interaction takes places within designated squads, found a significant effect of peers on academic GPA, especially during the first year. This supports Newcomb (1962) and Pascarella and Terenzini (2005) assertion that propinquity is an important peer-effect factor in post-secondary education.

Studies also found that low-achieving students were harmed through their placement with high-achieving students (termed 'Invidious Comparison' by Sacerdote (2011)). Meanwhile, average and high-achieving students benefited from placement with high achieving peers (termed “Boutique/Tracking” by Sacerdote (2011)) (J. Angrist, Lang, & Oreopoulos, 2009; Carrell, Sacerdote, & West, 2013; Feld & Zölitz, 2014; Garlick, 2013; Sacerdote, 2001, 2011).

Sacerdote’s (2011) review of previous peer-effect studies suggests that the negative effect of college-peers is documented, especially in residential colleges: Students who are assigned roommates who have a history of binge alcohol consumption in high school were more likely to engage in binge drinking in college (Duncan, Boisjoly, Kremer, Levey, & Eccles, 2005). In particular, the peer effect of fraternity membership on frequency of drinking was significant and increased the frequency of drinking (up to 20 percentage points) (Desimone, 2007). Other studies have documented peer-effect on various other negative behaviors, such as cigarette smoking (Wilson, 2007), marijuana use, and sexual activity (Duncan et al., 2005). Those peer effects, where one or more disruptive peer creates negative effects on surrounding fellows is referred to as the 'bad apple' effect (Sacerdote, 2011; Schuck & Zeckhauser, 2006 as cited in Bahr (2010)) .

Expanding upon previous research, Stinebrickner and Stinebrickner (2006) provided interesting insights to reveal the greater role played by curricular interaction over those of dormitory roommate interaction in the development of a student's social integration and establishing a peer network: This curricular interaction was demonstrated to be greater in terms of peer effect than those from roommate and dormitory interaction. Perhaps this finding relates to the positive peer effects learning communities have been found to have in studies on students at non-residential institutions (Chesebro, Snider, Venable, Green, & Mino, 1999, June; Engstrom & Tinto, 2008; Shapiro & Levine, 1999; Stassen, 2003; Vincent Tinto, 2003; Zhao & Kuh, 2004). In addition, they report that the effect of peers, while recognized, has minor influence over other factors including "the effort a student puts into studying, the quality of his/her study time, and his/her beliefs about the importance of educational attainment" (Stinebrickner & Stinebrickner, 2006, p. 1452)

The Role of Learning Communities.

In an effort to both replicate the cohort model and create synergistic linkages within the curriculum, learning communities encourage small groups of students to enroll in two or more connected courses. Students enrolled in a learning communities are more likely to have higher GPAs, earn more credit hours, and persist towards graduation at higher rates (Taylor, 2003). These higher rates of academic achievement are generally consistent regardless of institutional type, student demographic, or learning community format (Engstrom & Tinto, 2008; Knight, 2003; Price, 2005; Stassen, 2003; Vincent Tinto, 2003; Zhao & Kuh, 2004). The four most common formats of learning communities are: 1) paired or clustered courses; 2) cohorts in large courses; 3) team-taught programs; and 4) residence-based programs (Shapiro & Levine, 1999).

Vincent Tinto (1993) established that college persistence was tied to a student's ability to integrate within a college setting, claiming that students who developed strong ties (a sense of belonging) to the college were more likely to persist. Such connection is made possible through the creation of strong relations with peers, faculty and staff. In relation to learning communities, Vincent Tinto (2003) found that students in learning communities formed self-supporting groups; in class, and beyond the classroom setting. The learning community setting increases students' time spent together (inside and outside of the class) and increases time spent on learning activities. In addition, students in the learning communities are reported to be more active in the classroom. Vincent Tinto (2003) highlights the role of learning communities in creating not only "shared knowledge" (shared, coherent curriculum), but also "shared knowing" (enabling students to be familiar with each other) and "shared responsibility" (such that students are mutually dependent on one another so that the learning of the group does not advance without each member doing her or his part). It was found that in the context of a non-residential institution (i.e. community colleges) learning communities enhance academic outcomes. The longevity of the effect and its magnitude may vary. Scrivener suggested that the influence of peers in learning communities may concentrate in the semester in which students enrolled in the learning communities (Scrivener et al., 2008).

Limitations of Previous Studies- The challenge of using random sampling to measure peer effects, and alternative methods employing student-level *transactional data*

The analysis of peer-effects conducted-to-date suffered from the following deficiencies:

first- these studies are limited to trials in selective and highly selective institutions: two of the studies took place at Dartmouth College (Sacerdote (2001)); Williams College, (Zimmerman, 2003), both highly selective institutions; randomized trials of peer effects took place at West Point, (Lyle (2007)); and U.S. Air Force Academy, (Carrell et al. (2009)), all highly selective institutions in which randomized peer trials were conducted. Even a low-ranked student in a cohort at one of these schools is likely to be highly qualified from an academic perspective. Stinebrickner and Stinebrickner (2006) is among the few random assignment peer-effect studies that used students of disadvantaged background: However, because of the unique funding structure of the school, where all students receive full scholarship and housing, irrespective of family income, the admission pool is competitive (the admission rate for the institution tested is approximately 35 percent, making it a moderately selective institution). The selectiveness setting impacts the generalizability of the results to wider groups of postsecondary institutions. Randomized trials may provide a good estimation of the isolated 'randomized' peer effect, but are less applicable to the commuter and two-year college settings, where students have limited interaction time in learning activities outside the classroom setting (Center for Community

College Student Engagement, 2012; V. Tinto & Russo, 1994; Townsend & Wilson, 2008; Visher et al., 2010), or socializing through extra-curricular activities.

Second- previous studies assumed that students refer to their assigned group as peers, and assumed interaction of peers but did not test the assumption. In fact, previous studies noted that the estimated peer effects are sensitive to the definition of peers (e.g. in class, in dorms, in close proximity) and peer groups (Foster, 2006), as well as to the ability to measure peer's characteristics (Stinebrickner & Stinebrickner, 2006). In randomized trials evaluating peer effects (e.g. Sacerdote (2001); Lyle (2007); Carrell et al. (2009); Duflo, Pascaline, and Kremer (2011); Carrell et al. (2013); Feld and Zölitz (2014)) researchers based their findings on the assumption that an interaction between peers had indeed taken place. However, they did not measure the actual type or intensity of exposure to peers, either in-class or in other settings. The methodology of random assignment could not account for 'social' peers (Arcidiacono, Foster, Goodpaster, & Kinsler, 2012; Foster, 2006), and studies that employed random assignment provided limited, if any, measure of rate of exposure to peers (time or courses, number of peers, or proportion of the activities).

Third, previous studies on peer relationships represented either a semester or an academic year by taking a sample at a single point in time. This prevented the study’s ability to evaluate longitudinal differences in the exposure to peers or changes in an individual’s network of peers. In fact, and peers may move in and move out of an individual’s student network the exposure rate will change over time.

Fourth , Several scholars (Manski (1993), Sacerdote (2001, 2011), and Hanushek, Kain, Markman, and Rivkin (2003)) have noted a challenge evaluating causal relationships in peer-effect studies in terms of the pre-determined characteristics (unobserved variables), and in separating peer-effects from other covariates. Previous studies had a very limited set of covariates available and consequently the ability to control for unobserved variables was limited.

Student-level Transactional data allows to overcome those deficiencies, and provides important advantages to the current study:

First, the student-level transactional data analysis derives student-peer relationship information directly from course registration records. It is efficient at compiling complete data sets, repeatedly at key points in time, over complete populations. For example, in this study each student (N=3,128) was reevaluated every semester for academic performance, financial attitude, demographics, registration behavior leading up to the semester, peer associations, late registration, and class attendance. The key benefit being that additional important measurements of peers can be derived to better estimate peer-association and determine: How many peers does a student have, what portion of a student's classmates are peers, what is the exposure rate (repeated peer association, and longevity of peer associations) for student-peers. Previous studies which did not employ transactional, time stamped data, were far less effective or incapable, of deriving this information. Transactional data further permits to employ procedures in order to retrieve and post process records and to generate the summary information from administrative records.

Second, student-level student transactional data is useful to conduct peer association studies in settings where a randomized trial is not feasible. In this study, student-level transactional data allowed for the identification of and measurement of peer creation, and the impact of those peer relationships, all in the natural setting of a commuter, open admission community college. Transactional data allows for the measurements of student interaction to be taken indirectly, so that student peer association can be recorded even if the student is not available to complete a detailed survey or direct data collection regarding their experience in the college.

Third, student-level transactional data enables measurement of peer association and its influence overtime. It enables the evaluation of peer-interaction across semesters, without the need for follow-up surveys, and not only at a single point in time. It further enhances the quality of peer studies by its ability to control for the length and rate of exposure to peers, and by controlling for the exposure to 'recent' peers versus 'historical' peers.

Fourth, time stamped student-level transactional data allows researchers to measure the academic qualities of peers directly: The peer groups being identified through the administrative records, and summary information (i.e. academic standing) regarding individual peer qualities may be derived. It may provide greater accuracy of the student peers' quality compared with peer studies that do not employ an actual, direct measurement of students' association from existing records, and simply use regional or cohort summary profiles to assess peer-group impact.

Fifth, as the information available to community colleges is growing, using student-level transactional files allows colleges to revisit peer-association evaluations by adding new variables that are determined to be important or become available to colleges and universities (e.g. income data derived from tax records/FAFSA forms). Studies that do not use the information available at the transaction levels are limited in their ability to update and revisit their findings.

Lastly, employing student-level transactional data provides great accuracy in the student records in terms of an actual proximity to specific peers that cannot be achieved through other methods of data collection, in particular surveys and cohort estimation; it provides greater coverage (virtually for all participants) of their exposure to peers irrespective of response rate to surveys.

Chapter III

METHODS AND PROCEDURES

Sample.

Using data from a cohort of N=3,128 students entering a single, large, urban, public two-year community college, this study investigates the effect of student-level behaviors on their academic performance and progress in the degree. This cohort consisted of every student, first-time and transfers, whether full-time or part-time, which first entered the college at the beginning of the study period. The study aggregates student-level records for the cohort over a period of about six years (September 2003 through December 2009); purposely coinciding with the National Center for Education Statistics (NCES) Beginning Postsecondary Studies (BPS 2004:09) sample. This study demonstrates that measurements of additional, pattern/behavioral indicators can be retrieved from transactional, time stamped records, and can used by community colleges to better understand the combined risks influencing student pathways and affecting student outcomes. In addition, this study is useful in that it presents a methodology to identify useful behavioral indicators hidden within a comprehensive set of administrative student-level records and transactional data, circumventing the need to perform extensive intrusive data collection. Furthermore, by examining the joint effect of these indicators (late registration, schedule changes, peer effects and freshman seminars) this study enables the measurement of the outcome of these effects in the natural setting of one large, urban, community college; and comparing relative outcomes with those of students in more than 140 similar, community colleges nationwide (Title IV, urban, degree granting, with enrollment above 5,000 students, and 50 percent or more of the students receiving need-based aid) (U.S. Department of Education, 2009a, 2009b). Setting the timeframe for the study with that of BPS2004:09 allows for additional analysis and insights on student-outcomes of an equivalent period for students in urban, public, 2-year community colleges.

Data Structure and Methodological Advantages.

Adopting methods used in other higher education longitudinal studies (DesJardins, Ahlburg, & McCall, 1999; Laura Horn, 2009; Singer & Willett, 1993, 2003) the dataset for this study is structured in a student-period format. A baseline record was established for each individual student, and subsequent modifications to that record were recorded in additional individual rows. A summary is prepared for each student as one observation (one row) per semester. Outcome variables are included for each dependent variable to indicate if they occurred in semester (time) *t*. The study period is equivalent to the one covered by NCES' BPS 2004:09 (Horn & Nevill, 2006; NCES, 2012a). Students in the study continue to have a period *t* observation until the end of the period of the study (If they remain through the end of the 2009 follow up period).

The advantage of such longitudinal analysis is that in each time period, the time period itself is included in the estimation. It enables maximum flexibility on the parameterization of the model. (Cook, Campbell, & Shadish, 2006; Scott & Kennedy, 2005; Shadish, Cook, & Campbell, 2002; Singer & Willett, 1993, 2003; Smart, 2010). As such, the model can estimate coefficients for time-invariant and time-varying predictors.

For every semester in the study period, the data collected is extensive. It consists of a full disclosure of the student demographic (including updated information for student age, dependency status, international student status, etc.), information regarding the student academic standing and progress in the degree (including completing of remedial education for a particular term, when applicable; credits earned and attempted for the targeted semester, and cumulatively;

major in the semester; major or degree program changes), the student financial composite (both for the semester and cumulatively, using semester-based financial aid information), and behavioral indicators status (for the targeted semester, and cumulative). Student-level key academic indicators are verified against the student transcript, and through the national student clearinghouse (NSC). Using administrative records provides greater data coverage than possible through survey data collection: the record provides an official verification of the student's in a particular academic, demographic or financial category. Furthermore, the administrative record is available even in cases where the students have left the college or was not available to complete a questionnaire. Employing administrative records and transcript analysis ensures similar standards for data retrieval to the ones set in NCES' BPS 2004:09 (L. Horn, 2009; Horn & Nevill, 2006; NCES, 2012a), and the process ensures that the data coverage meets the standards set by NCES' procedures (Flemming, 1992; NCES, 2012b).

Table 1 illustrates the data structure in a student-period format, such as the one employed in this study:

Table 1

Data Structure, by Student, by Semester (Student-period format).

ID	Semester	Age	Late Registration	Cumulative Late	Semester GPA	Semester Credits	Cumulative Credits	Hispanic?	…
1234	**1**	**29.56**	**0**	**0**	**.**	**0**	**0**	**0**	…
1234	**2**	**30.06**	**1**	**1**	**3.784**	**5**	**5**	**0**	…
1234	**3**	**30.56**	**0**	**1**	**3.623**	**4**	**9**	**0**	…
2345	**1**	**46.36**	**1**	**1**	**3.426**	**3**	**3**	**1**	…
2345	**2**	**46.86**	**1**	**2**	**3.555**	**6**	**9**	**1**	…

Student-level records are constantly updated for each semester and for each student. By lieu of example, if a student registers late to classes in a given semester, their record for the particular semester is updated, as well as the cumulative record to date. If a student registers late for more than one semester, both the semester-based record and the cumulative record will be updated to reflect the change. Demographic (i.e. age), financial (i.e. aid recipient), and academic variables (e.g. GPA for the semester and cumulative, credits attempted and earned) are updated to reflect the student status in each semester. Each student is reflected in the dataset for each semester. In this study, the total of 3,128 students followed through this dataset resulted in 12,715 records of student-period (per student per semester), or an average of 4.06 semesters per student.

This study and data design carries several additional advantages derived from the availability of a rich, student-level transactional dataset. It promotes the use of student-level transactions reflected in common community colleges' information systems, specifically using behavioral characteristics as indicators to improve community colleges' ability to monitor and support student-level outcomes. The current study updates student-level profiles constantly, demonstrating the ability of community colleges to monitor and update individual level data. It accounts for constant changes of the student's demographic, academic and financial profile, and updates those changes on a regular basis. In addition, transactional data allows accounting for changes in indicators of student behaviors (e.g. schedule changes), financial (e.g. changes in aid status), demographic (e.g. marital status), and academic (e.g. change of major) throughout their studies. Third, the data used in the analysis is derived from transactional datasets that are universally required to complete college-level reporting (e.g. IPEDS), and available virtually for every community college (irrespective of computing environment, software or hardware infrastructure). Fourth, this study enables an evaluation of the magnitude (effect size) of different behavioral indicators on student outcomes, such that community colleges will be able to further assess the effect of each indicator in the context of their settings. Table 2 reports the variables available per student and per semester

Table 2.

Student-Period Variables (part 1 of 2)

Variable	Description	Valid Values
Student-Time Period Identifiers		
Rnd	Student ID (Unique Randomized Number)	Unique value maintained for each student
Time	Semester since the beginning of the study (Calendar time)	1 through 13
Termnum	Number of semesters actively enrolled	1 through 13
Student Behavioral Indicators		
Later	Late Registration to the semester	1=late, 0=on time
Cumlater	Cumulative number of late registrations	cont. number, may be zero
initialwit~1	Schedule changes post 1st day of classes	1=change,0=on time
cuminitial~1	Cumulative schedule changes from 1st day of classes schedule	cont. number, may be zero
Peers1	number of peers per semester	cont. number, may be zero, other peer-variables available
CPeers1	Cumulative number of peers per semester.	cont. number, may be zero, other peer-variables available
Student Demographics		
Male	Male Student	1=male, 0=female
Age	Age as of the first day of the semester	cont. number
Race/Ethnicity		
wt	White (non Hispanic)	1=Yes, 0=No
as	Asian	1=Yes, 0=No
hs	Hispanic	1=Yes, 0=No
bl	Black (Non-Hispanic)	1=Yes, 0=No
oth	Other Race/ethnicity	1=Yes, 0=No
intstu1	International Student Visa Holder in the semester	1=Yes, 0=No
Student Academics		
Pt	Part Time for the semester	1=Yes, 0=No
Catt	Credits Attempted in the semester	cont. values, may be zero
Cearn	Credits Earned in the semester	cont. values, may be zero
cum_catt	Cumulative credits attempted	cont. values, may be zero
cum_cearn	Cumulative Credits Earned	cont. values, may be zero
Tgpa	Semester GPA	cont. number, may be zero or null (if no valid GPA)
endsem_cum~a	Cumulative GPA at the end of the semester	cont. number, may be zero or null (if no valid GPA)
tcum_gpa	Cumulative GPA at the beginning of the semester	cont. number, may be zero or null (if no valid GPA)
Ttotcredit	Total Credits as of first day of the semester	cont. number, may be zero
numcourses	Number of courses in the semester	cont., may be zero (non-course, e.g. candidacy)
cnumcourses	Cumulative, Number of courses to date	cont. number, may be zero

Table 2.

Student-Period Variables (cont., part 2 of 2)

Variable	Description	Valid Values
Student Academics (Cont.)		
Graduated	Graduation Indicator	1=Yes, 0=No
ret1	Retained from Previous Semester	1=Yes, 0=No
gor1	Graduated or Retained from Previous Semester	1=Yes, 0=No
Courses and Major		
majchg1	Changed major from a previous semester	1=Yes, 0=No
cum_majchg1	Cumulative number of major changes since initial enrollment	cont. number, may be zero
Degree Program Data		
d_aas	Associate of Applied Science (AAS) Degree	1=Yes, 0=No
d_as	Associate of Science (AS) degree	1=Yes, 0=No
d_aa	Associate of Arts (AA) degree	1=Yes, 0=No
d_cert	Certificate program	1=Yes, 0=No
Admission Data		
adminitial	Initial Admission (First time/Transfer in)	Categories
ged	Earned GED	1=Yes, 0=No
m1num	Need Math remedial coursework (In a given semester)	1=Yes, 0=No
w1num	Need English (Writing) remedial coursework (In a given semester, Number format)	1=Yes, 0=No
Other Academic Variables		
fsm	Took Freshman Seminar or workshop	1=Yes, 0=No
stopout_back	Returned from Stopout in the semester	1=Yes, 0=No
cum_stopou~k	Cumulative returns from stopout	cont. number, may be zero
Completion Data		
grdgpa	Graduation GPA (If graduated in semester)	Value (4.000 GPA Scale, min=2.000, max=4.000)
grdcret	Graduated with a certificate (indicator)	1=Yes, 0=No (Implies AA/AS/AAS degree)
grdmaj	Graduation Major category	Categories
Student Finance and Aid		
pelltap1	Pell or TAP recipient (in a given semester)	1=Yes, 0=No
cumpelltap1	Cumulative Pell or TAP recipient	Number of semesters as TAP/Pell recipient
Tself	Tuition Paid through Self Resources, Sem Level (Non Aid)	Cont. values
Aself	Any Tuition paid through Self Resources (Non-Aid)	1=Yes, 0=No
Cself	Cumulative Tuition Paid through Self Resources (Non Aid)	Cont. values
Caself	Cumulative Num of Semsters, Any Tuition paid through Self Resources (Non-Aid)	Cont. values
Taid	Tuition Paid through Aid Resources, Sem Level (Any Aid)	Cont. values
AAid	Any Tuition paid through Aid Resources, (Any Aid)	1=Yes, 0=No
Caid	Cumulative Tuition Paid through Aid Resources (Non Aid)	Cont. values
Caaid	Cumulative Num of Semesters, Any Tuition paid w/aid	Cont. values

Table 2 demonstrates the richness of information available within virtually every college student records systems, using post-processing procedures of student-level transactional data. All information in the dataset is derived from common college's registration and transactional files. Student and time dependent profiles are constantly maintained and updated with semester-based data. The variables used in this study are divided into student behavioral indicators; demographics, academics, and financial variables; and for ease of use includes additional variables derived from the data. By maintaining longitudinal data, patterns may surface related to student behavioral patterns, demographic, and financial standing. At any point in time, the student's cumulative view to date is available. Along with their academic progress, this information can provide an enhanced evaluation of student progress towards academic milestones.

Data gathering through student-level Transactional Files and Administrative Records versus Survey Questionnaire

Using student-level transactional data and administrative records to analyze students' behaviors (in lieu of student-level surveys) carries numerous advantages which include: substantially increased data accuracy (Czajka, 2013; Koenig, 2003), the ability to perform exhaustive updates, and a reduction of errors and incomplete responses resulting from survey fatigue. The primary benefit is that administrative data allows a complete census of all students at any given moment. In theory, there is full coverage of the targeted group (Calderwood & Lessof, 2009), allowing researchers to overcome problems associated with low response rates, response bias, and survey fatigue. Secondly, administrative records usually go through routine maintenance cycles that allow for data 'cleaning', contributing to increased data accuracy, validation, and to future cross-referencing with other databases (Calderwood & Lessof, 2009;

Winkler, 2001). Robins, Rotnitzky, and Zhao (1994), Peytchev (2013), and others, demonstrated that the effect of non-response in survey instruments is not missing at random (NMAR), and therefore creates a bias on the estimators that rely on individual (student-level) data collection. Massey and Tourangeau report that survey fatigue and non-response rate has become a growing challenge for data collection procedures and for the statistical analyses across multiple governmental and non-governmental data collection efforts (Massey & Tourangeau, 2013b). Longstanding national surveys, such as the Survey of Consumer Attitudes, have reported a marked decline in response rates; from over 70 percent in 1979 to below 40 percent in 2005 (Massey & Tourangeau, 2013b).

Survey fatigue occurs when a particular group is presented with multiple surveys or excessively long surveys over a relatively short period of time (Adams & Umbach, 2012; Porter, Whitcomb, & Weitzer, 2004; Sarraf & Tukibayeva, 2014; Schuh, 2011). Survey fatigue has become an increasing concern as the cost of designing and administering surveys declines, and the access to technical tools that enable survey administration (in particular, web-based tools) increases. When presented with an ever-growing number of surveys, students are more likely to complete a survey partially or with perfunctory responses, or decide not to participate altogether (Dey, 1997; Porter & Whitcomb, 2005; Sax, Gilmartin, & Bryant, 2003).

Non-response, and partial response to student-level surveys has been a growing concern in higher education: Porter, Whitcomb and Umbach conducted a series of studies to evaluate various manifestations of survey fatigue and non-response analyses in colleges and universities (Porter & Umbach, 2006; Porter & Whitcomb, 2005; Porter et al., 2004). Their work suggests both a decline in survey response rates overall, and the prevalence of lower response rates at institutions located in urban areas. In addition, they report that a college's selectivity also affects response rates, such that colleges which have less selective admissions standards can expect lower response rates and lower survey completion rates. At the individual student level, students with higher levels of academic preparation and higher levels of attachment to an institution are more likely to complete surveys. These differences are responsible for creating a bias across different types of institutions and within open admission institutions such as community colleges this has a pronounced effect on survey coverage. Other findings suggested students responding to surveys were more likely to be females, socially engaged in the school, less likely to receive need-based financial aid, and more likely to be an investigative personality type (Porter & Umbach, 2006).

Methodology

The analysis employs both typical and transactional level data growth curve models in evaluating student outcomes. Growth curve models are part of a large set of mixed-effects modeling, which allows for fixed effects (intercepts and slopes meant to describe the sample as a whole, in this case- the students entering CC); and also for random effects, intercepts and slopes that can vary across subgroups of the sample. Unlike traditional mixed effects models, it allows for changes not only across subgroups or subjects, but also across time. Hamilton (2012) notes that mixed-effects modeling enables a range of possibilities for “multilevel o models, growth curve analysis, and panel data or cross-sectional time series” (p. 387). Models that employ student-level transactional based dataset and growth curve models are organized in a structure, such that

$$Y_{it} = X_{it}\beta + Z_{it}\delta_i + \varepsilon_{it}$$

with

$$\delta_i \sim N(0, G), \varepsilon_{it} \sim N(0, R)$$

Where δ capture systematic, *between* individual differences and the ε capture all *within* subject, unexplained differences. Subjects' data is available for each subject (i), and for each point in time (t). In typical models (i.e. logistic regression), these terms are univariate, independent of each other, with constant variance. But, if one accounts for the growth of information regarding the subjects (in this study, the students enrolled in community colleges) over time, then the differences between individuals are in fact a function of

$$G = \begin{pmatrix} \sigma_{\delta_0}^2 & 0 \\ 0 & \sigma_{\delta_T}^2 \end{pmatrix}, R = \sigma_\varepsilon^2$$

In such case, the X terms are potentially time-dependent predictors, Z is captures and individual-specific level and linear trend, by setting it as follows: Z =(1,time) for a single subject in the models.

The subject-specific (individual, student-level) effects $\delta=(\delta_0, \delta_T)'$ models the correlation structure within subject (one would expect individuals to maintain approximately the same level and trend, net of any other predictors/trends, across their enrollment period). But, one cannot pool time periods, and thus use all of the information available, without imposing some control for between subject differences as done in this study. *In other words, OLS applied to longitudinal data violates the independence assumption*. By employing these subject-level controls, one can ensure a form of robustness, as well, since time-constant differences between subjects are controlled (the explored fixed effects, rather than random effects model forms and the findings are robust to these two approaches to heterogeneity controls). Further robustness examination takes place through detailed analysis of the differences between the observed and expected values for student-level outcomes, both subject specific and over time.

Specifying Peers Exposure and Evaluating Peer Association

In this study, I set out to assess the measure of peer exposure directly from the students' enrollment record. The measure of peer exposure is defined as the number of a selected 'target' student's classmates who are repeatedly (twice or more) enrolled in the same classes as the target student. The student's peers are any students who enroll concurrently with the target student. Peer students are not only students in the study group cohort - Any student enrolled at the sampled institution, whether or not in the study cohort, could potentially be a peer, and the exposure of the targeted student to that peer is then recorded. I further test the measure of peer association by examining a stricter definition of peer exposure as the number of student's classmates who are enrolled in classes with the student three times, and four or more times. For all measures, I also record the number of unique classmates and the number of classmates seen by the students twice (or more), three times (or more), etcetera, so that the impact of the number of peers may be evaluated as well. The measurements are repeated every semester, and a cumulative measure is calculated for each semester as well.

Defining Peer Exposure Rates (Number of Peers)

In order to compile measurements of peer exposure the enrollment-record for every student (i) course(s) *(c)* within a specified time period *(t)* (semester, or cumulative period) is examined. For all students *(i)* enrolled in course(s) *(c)*, the enrollment of all other students in the courses (any student who enrolls with the target student, whether or not they are part of the targeted analysis) is also recorded. The records then allows to account both for the target students' and for students who took courses with the target students *(i)* and are not part of the targeted sample.

The total enrollment (duplicated) in all courses is

$$D_{i,t} = \sum_{1}^{c} N_{i,t} - C_{i,t}$$

where $N_{i,t}$= total number of students in each course, and $C_{i,t}$ is the number of courses taken by student *i* in period t,. The number of unique classmates of student i, over period *t* is represented by $U_{i,t}$, $U_{i,t} \leq D_{i,t}$. The number of peers, classmates who took more than one class with a student *i* in a specified timeframe *t* is represented by $R_{i,t}$ The number of seats taken by peer classmates is $D_{i,t} - U_{i,t}$, ($R_{i,t} \leq D_{i,t} - U_{i,t}$). For example, if a student takes four courses within a specified time-period ($C_{i,t}$=4), and if each course had thirty students enrolled (N=30), then the number of classmates for student *i* is $\sum_{1}^{c} N_{i,t} - C_{i,t}$=116 (120 seats minus the 4 seats taken by student i). To determine the percentage of peers over period (t), the number of students needed to fill in the seats the remaining seats (116), as well as the number of times each classmate was taking courses with the target student are recorded. In this scenario, if all 116 fellow students are unique (no repeaters), then the target-student has no exposure to peers, or $R_{i,t}$=0 .Conversely, if student (i) took all four classes (c=4) with all the same classmates, then s/he would have 29 peers, or $R_{i,t}$=29

Naturally, peer expose will vary as classmates may take courses with the target student once, or multiple times. For illustration purposes, let's assume over time (t) the 116 seats are taken by 99 unique classmates. The target student meets 89 classmates only once (unduplicated), and there are 10 other classmates (duplicated) that are met twice or more. If the combined student records show that two classmates took all classes with the target student (four classes); three classmates took three out of four classes with the target student; and the remaining five classmates took two classes with the target student, then the target student will have $R_{i,t}$=10 peers. To determine the percentage of peers that met with the target student *three times or more* then the number of peers would be $R_{i,t}$=5

Calculating the student peer exposer requires that for every student *i* and over each time frame *t*, the number of peers each student is exposed to is recorded. These values are then used to determine whether or not, and at what level, peer exposure can be correlated with changes in student outcomes, such as: credits earned, GPA, retention and graduation.

Time frames for Peer-Exposure Measurements: Semester and Cumulative Exposures

This study examines peer-exposure within a semester (point-in-time exposure to peers), and across semesters (cumulative exposure to peers). In the natural setting of a commuter community college, students may take any number of courses in a particular semester. In addition, most community colleges do not establish a cohort framework for students to follow. Therefore, peer exposure may present itself immediately, or develop over time. A student may take a single course in a specific semester, yet establish a small network of peers in that course and maintain enrollment in sections that have the same peers in subsequent semesters. For that reason, I measure the exposure of students to peers (peer association) not only within the semester, but also across semesters. The study employs the following calculation of cumulative exposure to peers:

For each student in each semester, the cumulative aggregate of peer exposure to-date is calculated. All measures of peer exposure is carried forward through all subsequent calculations. For example, if the student enrolled for a total of five semesters to date, the enrollment records from the first through the fifth semester are taken and evaluated for peer exposure across all five semesters. The measures of cumulative exposure are stored for each of the five semesters, such that $R_{i,kt}$, $U_{i,kt}$ the number of peers and the number of unique peers (respectively) is recorded; In this notation *kt* stands for cumulative peer exposure at point-in-time *t* (e.g. k5 would be cumulative through the fifth semester), $R_{i,kt}$ stands for the number of classmates who were peers with a student *i* at any point within timeframe *kt*; and $U_{i,kt}$ stands for the number of unique classmates student *i* has had at any point during timeframe *kt;* and $C_{i,kt}$ stands for the number of courses student *i* enrolled in during the cumulative timeframe *kt.* This process is repeated to generate a cumulative record at each point in time prior to the last semester the student enrolled, and stores it to the data file

Estimating the Association of Peer Exposure and Student Outcomes

Growth curve models used in the evaluation of peers' influence on student outcomes are part of a large set of mixed-effects modeling, which allows for fixed effects (intercepts and slopes meant to describe the sample as a whole, in this case- the students entering CC); and also for random effects, intercepts and slopes that can vary across subgroups of the sample. Unlike traditional mixed effects models, it allows for changes not only across subgroups or subjects, but also across time (Singer & Willett, 1993, 2003). At each point in time *t,* the study evaluates the association between exposure to peers and student-level outcomes, specifically-student's semester GPA, cumulative GPA, and credits earned in a growth model represented by

$$Y_{it}=P_{it}\beta+ Z_{it}\boldsymbol{\delta}_i+ Z_{it}\boldsymbol{\gamma}_i+ Z_{it}\eta_i+ \varepsilon_{it}$$

Where β captures the influence of peers exposure (P_{it}) for student *i* at time *t* . This impact is evaluated both as exposure rate (percentage) and using the number of student-peers*;* $\boldsymbol{\delta}_i$ $\boldsymbol{\gamma}_i$, and $\boldsymbol{\eta}_i$ capture systematic differences between individual in their demographic, academic, and financial profile. Specifically, $\boldsymbol{\delta}_{i,}$ represents the demographic profile (age as of the first day of the semester, gender, race/ethnicity profile, and international student visa status); $\boldsymbol{\gamma}_i$ accounts for differences in the students' academic profile (remedial preparation status as of the first day of the semester in mathematics, remedial preparation status as of the first day of the semester in English, credits earned as of the first date of semester, GPA as of the first date of semester, FT or PT enrollment status for the semester, number of classes attempted within the current semester, and GED recipient status) and $\boldsymbol{\eta}_i$ accounts for the financial profile (amount and percentage tuition paid through any aid resource, amount and percentage paid through self-resources, and in addition also track TAP or Pell recipient indicator status for the semester). ε_{it} captures all *within* subject unexplained differences. Subjects' data is maintained for each subject *i*, and for each point in time *t*. The estimated outcomes for each model are stored and used to further evaluate the quality of the model.

The models are be estimated by adding variables in a block order. First, the models are estimated using peer-associations only. Then, the demographic variables block is added to the model. Third, the academic variables block is added, and lastly - the financial indicators block is added to the model. Testing the model in this order allows for the evaluation of peer association in the presence of additional covariates. It also allows controls for alternative explanations to student peer-interaction that may be reflected in other variables.

Causality and Association in the Estimation of Behavioral Indicators' Influence on Student Outcomes.

Previous studies of behavioral indicators have noted that there is a challenge identifying causal relationships to peer-effects because of pre-determined characteristics (unobserved variables), and with regard to separating peer-effects from other confounding covariates (e.g.

Manski (1993), Sacerdote (2001, 2011), and Hanushek et al. (2003)). The challenge stems from an inability to control for covariates that pre-date the interaction, and also with the studies' inability to control for unobserved student characteristics and environmental qualities that may have also influenced student outcomes. The common approach used to overcome this challenge was through random assignment of students into peer groups. Random assignment assumes that unknown differences will not be correlated with the assignment of subjects into the control or experimental groups.

This study does not use randomization in its design, and did not control student exposure to peer associations. A randomized trial requires a prospective setting and typically involves a limited sample from the target population. Through the use of the student-level transactional data, the ex post facto design of this study identified naturally occurring peer relationships directly through observation of an entire community college entering cohort (N=3,128) over the entire study period (13 semesters). The advantage over previous studies that evaluated peer association using student or administrative surveys (e.g. (Crosnoe, Riegle‐Crumb, Field, Frank, & Muller, 2008; Riegle-Crumb, Farkas, & Muller, 2006)) is the inherent accuracy of EDM which permits measurement of the actual proximity among specific peers, not possible through other methods of data collection, including randomization, and certainly not to the scale achieved in this study. This study is able to provide a precise descriptive of the student peer association at any given point in time or cumulatively throughout the students enrollment. At the same time, this study may be limited by concerns regarding selection on unobservable variables. When this concern was raised in previous studies its focus was centered on the selection of courses (type) which may differentiate the student's (and peers) abilities. (Arcidiacono et al., 2012). The current study is not as likely to be influenced by this limitation for two reasons:

First, the students' foundation courses and remedial coursework are offered in multiple course sections (some were offered in as many as 40 different sections each semester). Students are unlikely to self-select (as a group) to concentrate in a single course section. Even if students would concentrate at high-demand times during the week (e.g. mid-morning, midweek classes), they were presented with multiple, concurrent sections of the same foundation courses.

Second, the construction of the regression models in this study take into account a rich set of academic, demographic and financial covariates. It accounts for numerous covariates unavailable to other studies due to data collection, specification, granularity, and survey size limitations.

When compared to previous studies- this study accounted for academic, demographic and financial variables at baseline, and performed updates at key points in time each semester to reflect changes over time. Those measurements accounted for common unobserved variables that may influence student outcomes reflected through peer association. While randomization does not take place, through an exhaustive revisit of the covariates each semester, this study provides enhanced control to the 'order of time' influence. In other words- covariates are measured at the beginning, during, and at the end of each semester allowing to better evaluate the influence of student peer association in the semester, and cumulatively, while controlling for other known covariates . These controls account for the students' academic, demographic and financial composite immediately prior to the measurement of the behavior indicators (*in each semester*) and their association with the students' outcomes. Such controls follow the recommendations of Schneider, Carnoy, Kilpatrick, Schmidt, and Shavelson (2007) to correct for any potential selection bias in the establishment of peer association. Schneider, et al, noted that "correcting for selection bias when there are identifiable treatment groups is to adjust for fixed, unobserved characteristics that may be associated with selection into the treatment group" (p. 42). This study could not eliminate the possibility that some unobserved student characteristics were reflected in behavioral indicators. In response, by controlling for a rich set of pre-existing qualities of each student (demographic, academic and financial) at numerous points in time minimizes the potential influence of unobserved characteristics that may be reflected through behavior indicators. Furthermore, this study provides a detailed student behavioral descriptives over a longitudinal period which was not available in previous studies of community college student cohorts. While the statistical procedures in this study, similar to other studies (Schneider et al. (2007), pp. 95) reduce the bias in the analysis, it is important to note that possibility that observed or unobserved characteristics that are associated with student academic outcomes (specifically, GPA, credits earned and retention or graduation) is omitted from the analyses. Following this concern, this study can demonstrate associations between student peer networks, late registration, schedule changes and the observed academic outcomes; but one should use caution in determining causal relations or using causal language in the interpretation of the results. Nonetheless, this study provides an important layer to the literature by demonstrating the existence of student peer association. In addition, this study measures student peer associations where peers cannot be assigned randomly, provide an enhance mapping of the establishment (creating and dissolving) peer association network, and evaluates the relationships between student behaviors and peer association, and academic outcomes, which can be in future studies explored in experimental settings.

Models Quality and Robustness Estimation

The estimation of the models' stability is conducted using both in-sample and out-of-sample estimation. Correlation between predicted and actual outcomes for in-sample estimations are calculated for each point in time (e.g. semester), and also cumulatively (for entire study period). These estimations allow the study to determine if the prediction quality of the model changes (ideally improves) over time. A model providing stable estimation will correlate well between estimated and actual outcomes. Out-of-sample estimation is obtained using random subsets of the records that can be used to predict the outcomes of all remaining students in the dataset. The analysis further estimates the correlations between the model's predicted outcomes using a subset of the results and the outcomes using all remaining records. A stable model will result in strong correlations between out-of-sample prediction and the full sample predictions. While there is no single standard for out-of-sample testing, it is common to generate subsets of 20 to 25 percent of the cases for out-of-sample models' testing, and this study follows the same practice.

Software Solutions and IRB

All models are analyzed using Stata (StataCorp., 2014). The unconditional growth models are estimated using *xtmixed/mixed* with both random intercepts and random slopes for time. Binary (logistic) dependent variables models are estimated using *xtmelogit* and *meqrlogit*. Predicted estimations are then stored using the *predict* command and are further analyzed for robustness. Models specifications follow the recommendations of Albright and Marinova (2010), and Hamilton (2012). The study and the research protocol were reviewed and approved by the Institutional Review Board (IRB).

CHAPTER IV

STUDENT PEER ASSOCIATION

This study captured thirteen concurrent semesters of student transaction data, to describe curricular peer association among students. For the purposes of this study, peer-association is defined as the coincidence of any two students enrolling together in more than one course during the study period. The rate of peer-association among students is derived by identifying coincidental course enrollment each semester, and then determining to what extent a pattern develops between individuals over time. Identifying patterns is based on the rate of peer-association for individual students, which counts the number of times an individual student enrolls with another unique individual student. Peer students are any students who take a class concurrently with the sampled student. Aside from no association at all (zero) this rate will range from"1", a pair of individuals having an association only twice during the study period (perhaps during a single semester), to an individual maintaining associations with multiple peers over multiple semesters. There is no attempt to understand the motivation of why any two students enrolled in the same course. The interest is limited to the fact that two students enrolled in the same course; and that their enrollment persisted, at least, from the first day of class through to the census date. Ultimately, the goal is to discern if peer-association results in peer-effects statistically correlated to student academic outcomes ($H1_c$: Evidence of peer-effect is associated with improved academic performance).

First Semester Peer Association

First step is to evaluate the study group cohort in their first semester's enrollment by describing the population as related to observable peer-associations. The study cohort, N=3,128 students, are part of a total college student population of 17,744. The cohort is made up of first-time college students and new transfer students. Any student at the tested college continuing from the previous semester, returning from an unofficial stop-out, or readmitted was excluded from the study cohort. Consequently, students in the study's cohort enrolled in courses with other students from the general student body, and therefore peer associations are counted among members of the study cohort and anyone else attending the college. Enrollment in remedial education and entry-level first-year courses is very common of both the general student body and the study cohort: Intermingling of these two groups is inevitable.

The peer associations were made according to 12 categories (Number of Peers within the entire college student body, Age, Gender, Race/ethnicity, Academic preparation, New/transfer, Late registration, Schedule changes, Participation in new student seminar, number of classes taken, GPA, and Persistence). What will be seen of the study cohort in the first semester is that the rate of peer association varies widely among individual students, and that on average a student's repeated exposure to the same peer(s) (two or more classes) is low (5.4%). As one would expect, those students enrolled in the most courses are more likely to have peer-associations within a given semester, and overtime. Also, students in the study cohort who participated in new student seminar had more peer associations. Students who were less likely to have peer associations included older students (greater than 25 years old), transfer students, and international students. In addition, non-remedial students, late registrants and schedule changers were less likely to have peer associations.

Entire cohort peer association.

The average student in the study cohort enrolled with approximately 107 unique classmates among the entire student population during their first semester (based on all courses taken). Table 3 shows that of these classmates, the average peer association was quite low, only 5.8 unique students, or 5.4% of classmates.

Table 3

Average Number of Unique Classmates for Each Student in the First Semester, and the Percentage of Peer Associations

Semester	Unique Classmates	Unique Peers (Two or more classes)	Unique Peers (Three or more classes)	Unique Peers (Four or more classes)
1	106.8	5.8 (5.4%)	3.2 (3.0%)	2.5 (2.3%)

The average student also met some classmates three or more times in the first semester. Within the peer associations, 3.2 (or 3.0%) where enrolled together three times or more, and 2.5 (or 2.3%) where enrolled together four times or more.

Student demographics and peer association.

Age

Table 4 examines the differences in age for students who have, and do not have, peer associations during their first semester, by the number of peers.

Table 4

Mean (SD) Age of Entering Students by Number of Peer Associations in the First Semester

	All Students	0	1	2	3	4	5+ Peers
Mean	24.2	26.8	24.1	23.8	23.4	23.0	23.5
SD	(7.6)	(9.2)	(7.4)	(6.8)	(6.9)	(6.3)	(7.2)

The mean age (SD) of entering students in the cohort is 24.2 (7.6). A student identified as having no peer associations is referred to as a '*soloist*.'[1] Soloists in the first semester were older than the mean ($p<0.05$), averaging 26.8 years old.

[1]Peer association occurs when a student is enrolled in the same courses with another student two or more times during the study period. By definition, a student enrolled in a single course will be a soloist. However, it is possible for a student to enroll in numerous courses and still be a soloist.

Students with four and five (or more) peer associations were the youngest ($p<0.05$), 23.0 and 23.5 years old respectively. The other students, those with between one and three peer associations were slightly younger than the mean age (23.4 and 24.1 years old), but that age difference was not statistically significant.

Race/Ethnicity

Table 5 reports the distribution of peer association by race/ethnicity group during the first semester.

Table 5

Proportion of Entering Students in each Race/Ethnicity Category and the Number of Peer Associations in the First Semester

# of Peer Associations	All (N=3,128)	AA (n=535)	Hispanic (n=1,044)	Asians (n=460)	Whites (n=413)	Int'l Students (n=367)
0	18.1%	17.8%	15.3%	19.8%	19.1%	27.2%
1	14.2%	12.9%	13.5%	14.1%	13.8%	15.3%
2	12.9%	13.5%	12.6%	12.8%	10.9%	9.3%
3	11.2%	14.0%	10.9%	10.0%	9.9%	11.2%
4	8.0%	9.7%	8.0%	5.7%	8.2%	8.7%
5+	35.6%	32.1%	39.6%	37.6%	38.0%	28.3%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

When compared to soloists in the entire cohort (18.1%), international students (student visa) were the most likely to be soloists during their first semester (27.2%). Otherwise, these results suggest that the distribution of students who are soloists does not vary much by the student's race/ethnicity profile. At the other end of the spectrum, students who encounter numerous peers, international students were also least likely to develop a large number of peer associations (5 or more) than the entire cohort, 28.3% and 35.6%, respectively. Taken together these differences for the international category appears to suggest they have the fewest peer associations.

From a different perspective, Table 6 reports the distribution of peer associations per capita by race/ethnicity category. Per capita in this context refers to the percentage of peer associations observed of each student in each race/ethnicity group.

Table 6

Distribution of Peer Associations Per Capita by Race/ethnicity in the First Semester

	% Peer Associations Per Capita	95 % Confidence Interval
All	5.4	5.16-5.67
Asian	6.8	6.03-7.49
Hispanic	6.0	5.56-6.48
AA	4.1	3.56-4.57
Whites	5.9	5.15-6.63
Int'l	5.5	4.76-6.20

($p<0.005$)

When viewed in this way Asians are seen as having the highest percentage of peer associations (6.8%) and African Americans have the lowest, peer associations per capita (4.1%). Using Bonferroni's post hoc test Table 7 reports the significant level of the differences in peer association per capita among race/ethnic groups

Table 7

Significant Level of Differences of Peer Associations Per Capita between Race/Ethnicity Groups throughout the Study Period

	Int'l	Asian	AA	Whites	Hispanics
Asian	n/s				
AA	0.047	0.000			
Whites	n/s	n/s	0.000		
Hispanics	n/s	n/s	0.021	n/s	

Bonferroni post hoc test

African American students had significantly lower peer association per capita than all other race/ethnic groups. The African American peer association level (4.1%) was 1.3% lower than the overall cohort (5.4%), and between 1.4% and 2.7% lower than other Race/Ethnic groups.

Gender

Table 8 reports the distribution of peer association by gender during the first semester.

Table 8

Proportion of Entering Students by Gender and the Number of Peer Associations in the First Semester

# of Peer Associations	All (N=3,128)	Males (n=1,201)	Females (n=1,927)
0	18.1%	17.5%	18.5%
1	14.2%	14.9%	13.8%
2	12.9%	12.4%	13.2%
3	11.2%	12.1%	10.7%
4	8.0%	7.7%	8.1%
5+	35.6%	35.5%	35.7%
Total	100.0%	100.0%	100.0%

There were no significant differences in the proportion of peers for male and female students in the first semester. Again, using the per capita perspective, Table 9 reports the distribution of peer associations per capita by gender.

Table 9

Distribution of Peer Associations Per Capita by Gender in the First Semester

	% Peer Associations Per Capita	95% Confidence Interval
All	5.4	5.16-5.67
Female	5.6	5.23-5.89
Male	5.2	4.79-5.59

There were no significant differences in the proportion of peers for male and female students in the first semester. Yet, when viewed as peer associations per capita, males have slightly lower peer association levels than females, 5.2% and 5.6%, respectively.

Student Academics and Peer Association.

Admission Status (first-time enrolment vs transfer)

First-time students have a greater number of peer associations in the first semester than transfer students. First-timers are more likely enroll according to prescribed course sequences, electing degree programs and services, such as learning communities, remedial clustered courses, and freshman seminar, which results in more peer associations. In this study cohort 74 percent of first-time students needed remedial course work compared to 41 percent for transfers. Similarly, 55 percent of first-timers took freshman seminar compared with 37 percent of transfer students. Table 10 evaluates differences in the proportion of students with peer associations by their initial admission status in the first semester.

Table 10

Proportion of Entering Students and the Number of Peer Associations by Admission Status in the First Semester.

# of Peer Associations	All Students (N=3,128)	Transfer-in (n=1,103)	First time students (n=2,025)
0	18.1%	27.9%	12.8%
1	14.2%	16.9%	12.7%
2	12.9%	13.7%	12.5%
3	11.2%	9.8%	12.0%
4	8.0%	6.1%	9.0%
5+	35.6%	25.7%	41.0%
Total	100.0%	100.0%	100.0%

(p<0.001)

First-time students, many of whom had 5 or more peer associations (41%), were more likely to establish multiple peer associations when compared to transfer students (25.7%). Conversely, transfer students were more likely to have one or less peer associations (44.8% versus 25.5%for first-timers). The differences in the proportion of students with peer association between first-time and transfer students were significant ($p<0.001$).

College Readiness *(academic preparation level)*

Table 11 reports the proportion of students with peers by their academic preparation level in the first semester. Students' college readiness is assessed by their high school and advance placement credits (or student transfer credits), Regents' test results, SAT or ACT scores, and/or student placement exams (COMPASS and ACT taken at the time of college admission in Mathematics and English).

Table 11

Proportion of Entering Students who are College-Level vs Remedial Coursework Needed by the Number of Peers in the First Semester.

# of Peer Associations	All Students (N=3,128)	College-Level (n=1,187)	Need Remedial Coursework (n=1,941)
0	18.1%	23.7%	14.7%
1	14.2%	14.6%	14.0%
2	12.9%	11.8%	13.6%
3	11.2%	8.3%	13.0%
4	8.0%	6.8%	8.7%
5+	35.6%	34.9%	36.0%
Total	100.0%	100.0%	100.0%

($p<0.05$)

Significant differences in peer-associations appeared between students needing remedial coursework versus those who tested as being college ready. Students who tested at college-level readiness (or exempted from remedial testing) were less likely to have peer associations in the first semester compared with students needing remedial coursework. This is to be expected, as entering students taking remedial coursework are likely to have their coursework clustered with additional (credit bearing or remedial) coursework. On the other hand, college-level foundation courses are a more heterogeneous mix of students from all disciplines and course sections are offered in great numbers, and widely dispersed throughout the week. As a result, college-level students are perhaps less likely to meet a peer multiple times during the semester.

Behavioral Indicators

On-Time vs. Late Registration

Having peers was found to be correlated with student-level behavioral indicators. Table 12 reports the proportion of students with peers by their registration pattern for the first semester (on-time or late registration).

Table 12

Proportion of Entering Students and the Number of Peers by Registration Pattern in the First Semester.

# of Peer Associations	All Students (N=3,128)	Late Registrants (n=412)	On-time registrants (n=2,716)
0	18.1%	28.9%	16.5%
1	14.2%	16.5%	13.8%
2	12.9%	10.4%	13.3%
3	11.2%	6.8%	11.9%
4	8.0%	6.1%	8.2%
5+	35.6%	31.3%	36.2%
Total	100.0%	100.0%	100.0%

(p<0.001)

Registration patterns are related to significant differences in the establishment of peer associations. Late registrants (those students registering on or after the first day of classes) are less likely to establish any peer association compared with on time registrants: 28.9 percent of the late registrants in the first semester are soloists, compared with only16.5 percent of on-time registrants. Similarly, 45.4 percent of late registrants had one or less peer associations (compared with 30.3 percent of on time registrants). As the semester begins, finding available seats in classes becomes difficult, so students who practice late registration are less capable of coordinating their registration with other students. Therefore, they are less likely to enroll in sections that would allow them to deliberately maintain peer associations.

Furthermore, students who register late differentiate not only in the number of peers, but also in the proportion of peer associations among their classmates. Figure 1 presents the relationship between the number of peer associations and percentage of peer associations for on time and late-registering students.

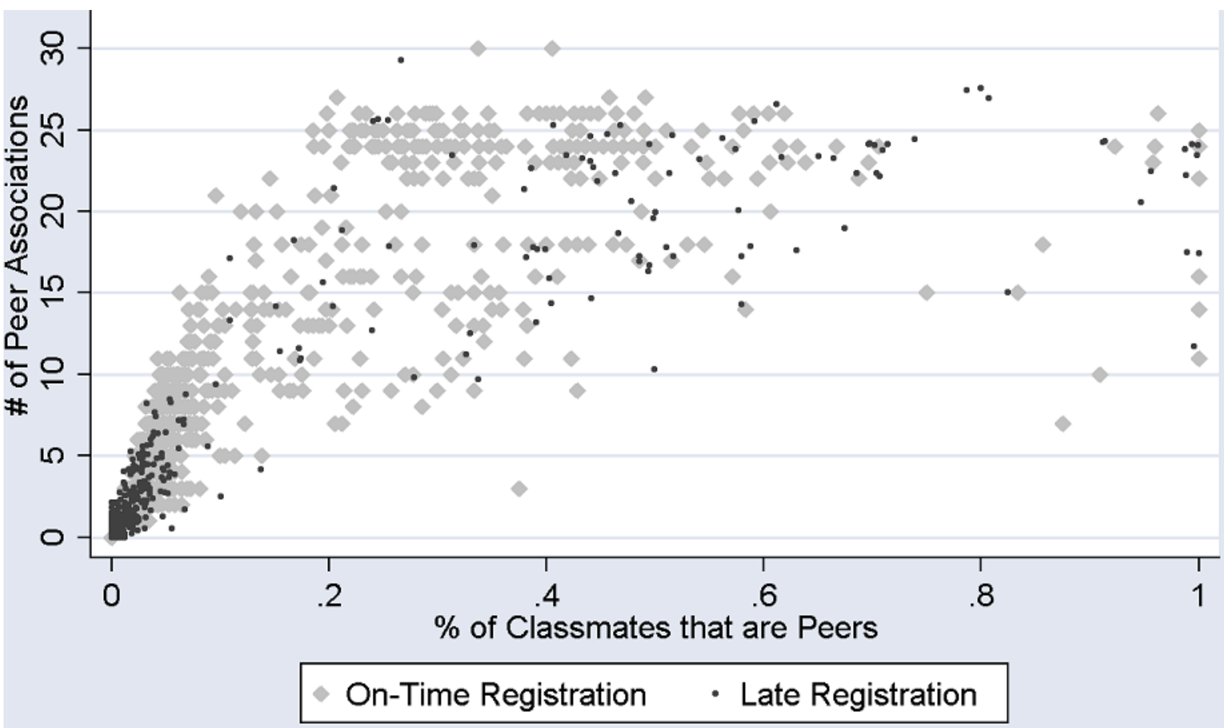

Figure 1. The Number and Percentage of Peer Association out of All Classmates for Students in the First Semester, by Registration Pattern

Figure 1 shows that late registrants have fewer peer associations in the first semester compared with on-time registrants. Late registrants displayed an interquartile range (IQR) of 0-6 peer associations compared with an IQR of 1-6 peer associations for on-time registrants. This means that more than 25% of late registrants are soloists (compared with 16.5% soloists among the on-time registrants). Late registrants had a median number of peer associations of only 2 compared with a median of 3 peer associations for on-time registrants. In addition to having fewer peer associations, late registrants' peer associations consisted of a smaller percentage of their classmates (Median 1.8% IQR 0-6.2%) versus a median of 2.4% (IQR 1.0-5.0%) for on-time registrants.

Student Initiated Schedule Change

Student initiated schedule change is related to lower peer association in the first semester. Table 13 compares peer associations of students who changed their initial course schedule with those students who retained their initial course schedule.

Table 13

Proportion of Students and the Number of Peer Associations by Pattern of Schedule Changes in the First Semester.

# of Peer Associations	All Students (N=3128)	Changed Initial Schedule (n=1291)	Retained Initial Schedule (n=1837)
0	18.1%	19.2%	17.4%
1	14.2%	17.0%	12.2%
2	12.9%	13.9%	12.2%
3	11.2%	11.4%	11.1%
4	8.0%	7.3%	8.4%
5+	35.6%	31.1%	38.7%
Total	100.0%	100.0%	100.0%

(p<0.001)

Students who retained their initial class schedule (approximately 59% of the study cohort) were more likely to establish peer associations (82.6 %) and more likely to establish multiple peer associations (70.4%) than in the case of schedule-changers, 80.6% and 63.8%, respectively. This difference was statistically significant. Among the students who changed their initial schedule, 19.2%were soloists in the first semester, and 36.3% ended up with one or less peer association. Among students who retained their schedule, only 17.4% were soloists, and 29.6%had one or less peers. In addition, students who retained their initial schedule were more likely to establish a peer association with five or more peers than students who changed their initial schedule, 38.7% and 31.1%, respectively.

Student Enrollment in New Student Orientation (Workshop or Seminar)

Student enrollment in new student orientation is related to higher peer association in the first semester. Table 14 compares the number of peer associations of students who enrolled in freshman seminar with those students who did not register for freshman seminar.

Table 14

Proportion of Students and the Number of Peer Associations by Freshman Seminar Enrollment the First Semester.

# of Peer Associations	All Students (N=3,128)	Enrolled in Freshman Seminar (n=1,606)	Not Enrolled in Freshman Seminar (n=1,522)
0	18.1%	8.5%	27.2%
1	14.2%	12.9%	15.4%
2	12.9%	13.1%	12.8%
3	11.2%	12.0%	10.5%
4	8.0%	9.8%	6.2%
5+	35.6%	43.7%	27.9%
Total	100.0%	100.0%	100.0%

($p<0.001$)

Students who enrolled in freshman seminar (approximately 51% of the study cohort, n=1,606) were more likely to establish peer associations (91.5 %) and more likely to establish multiple peer associations (78.6%) than students who skipped freshman seminar, 72.8% and 57.4%, respectively. This difference was statistically significant. Among the students who skipped freshman seminar (approximately 49% of the study cohort, n=1,522) 42.6% ended up with one or less peer association and 27.2% had no peer associations at all in the first semester. Among students who enrolled in freshman seminar in the first semester, 21.4% had one or less peers and only 8.5% had no peer association. In addition, students who enrolled in freshman seminar were more likely to establish a peer association with five or more peers (43.7%), compared with 27.9 % of students who had not enrolled in freshman seminar.

Student Course Load, First Semester GPA, Retention and Peer Association

Course Load.

During their first semester students enrolled in 4.6 (SD 1.6) courses on average, with variations ranging from a low of 1 course taken, to in rare cases, a high of 10 courses. Naturally, the number of courses a student enrolls in effects the chances of students meeting at random, and also opens opportunities for students to encounter each other more often. A student identified as having no peer associations are referred to as 'soloist.' As described earlier, a peer association occurs when a student is enrolled in the same courses with another student two or more times during a semester. By default, a student that enrolls in a single course will be a soloist. It is also possible for a student to enroll in numerous courses and still be a soloist. Table 15 examines peer association of students by the number of courses a student registered for in the first semester.

Table 15

Proportion of Students and the Number of Peers Associations by Course Load in the First Semester.

	Course Load					
# of Peer Associations	**1**	**2**	**3**	**4**	**5+**	**Total**
0	161 (100.0%)	114 (60.0%)	99 (32.1%)	98 (14.8%)	96 (5.3%)	567 (18.1%)
1	0 (0.0%)	44 (23.2%)	85 (27.6%)	138 (20.9%)	177 (9.8%)	444 (14.2%)
2	0 (0.0%)	14 (7.4%)	57 (18.5%)	126 (19.1%)	207 (11.4%)	404 (12.9%)
3	0 (0.0%)	8 (4.2%)	36 (11.7%)	104 (15.7%)	203 (11.2%)	351 (11.2%)
4	0 (0.0%)	1 (0.5%)	12 (3.9%)	53 (8.0%)	183 (10.1%)	249 (8.0%)
5+	0 (0.0%)	9 (4.7%)	19 (6.2%)	142 (21.5%)	942 (52.1%)	1,113 (35.6%)
Total Students	161 (100.0%)	190 (100.0%)	308 (100.0%)	661 (100.0%)	1,808 (100.0%)	3,128 (100.0%)

As the number of courses a student enrolls in increases, their likelihood of remaining a soloist decreases. Although 95% of students took more than one course in the first semester, 18.1% of the students were soloists. As the number of courses taken increases, the proportion of soloists' declines from 60% at two courses to 5.3% for students taking five or more courses.

A review of table 15 suggests that the number of peer associations does not increase at the same rate of the number of courses. Perhaps some of the peer associations may be a result of learning communities' course bundles, and some peer associations may be a result of chance alone. Figure 2 compares the proportion of soloist students (using a slightly conservative definition of peer association where students need to meet *three times or more* to be considered as having a peer association) and under the initial definition (where students need to meet only twice or more to be considered as having peer association).

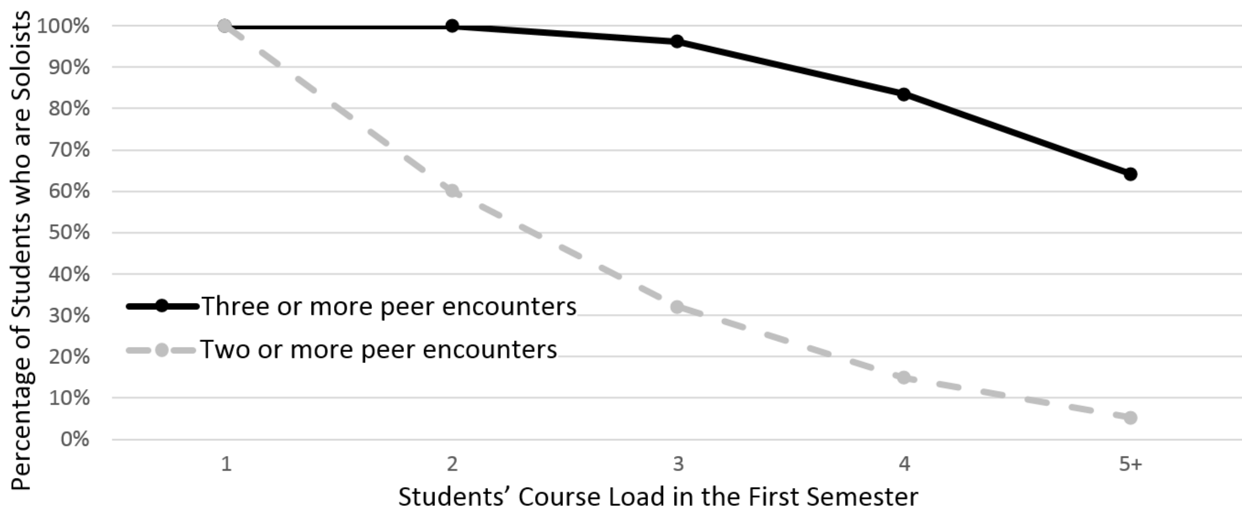

Figure 2. Percentage of Students who are Soloists with 2 or more, or 3 or more, Peer Association Encounters by Course Load in the First Semester

Overall, redefining the number of times peers meet to define establishing a peer association, from two to three, increased the proportion of soloists in the first semester from 18.8% to 75.4%. The decline in the proportion of soloists as the number of courses increases is notably smaller under this slightly stricter definition of the number of peers. At five or more courses, only 5.3% of the students are soloists when the definition of association is two or more encounters, compared with 64.2% of the students when the definition employed is three or more encounters. This change suggests that a singular peer association may be partially attributed to randomness and should be carefully examined. (A detailed review of the number of peers and the number of courses under a stricter definition in the first semester is available in appendix A)

GPA in the First Semester

Table 16 evaluates differences in student's first semester GPA by the number of peer associations established.

Table 16

First Semester GPA by the Number of Peer Associations.

# of Peer Associations	# of Students	Students in Credit-Bearing Courses	Mean GPA	SD
0	567	490	2.210	1.440
1	444	419	2.245	1.353
2	404	380	2.249	1.345
3	351	317	2.322	1.277
4	249	240	2.359	1.287
5+	1,113	1,094	2.495	1.219
Total	3,128	2,940	2.350	1.310

Note: Bold type indicates significance ($p<0.05$)

When compared with the entire cohort, soloists have earned a lower GPA (2.210 vs. 2.350, $p<0.05$). Conversely, students with numerous peer associations (5 or more) have earned a higher GPA compared with the entire cohort (2.495 vs. 2.35, $p<0.05$). Overall, as peer associations increased there was a trend of increasing GPA.

Retention to the Second Semester

One of the underlying hypotheses guiding a peer association analysis is that students with more peer associations would differ in their academic outcomes from those students who establish fewer peer associations. Table 17 evaluates differences in the proportion of students with peer associations and their retention into the second semester.

Table 17

Retention from First to Second Semester by Number of Peer Associations.

# of Peer Associations	All Students (N=3,128)	Did not return (n=745)	Retained to 2nd Semester (n=2,383)
0	18.1%	30.7%	14.2%
1	14.2%	14.1%	14.2%
2	12.9%	12.1%	13.2%
3	11.2%	10.2%	11.5%
4	8.0%	6.7%	8.4%
5＋	35.6%	26.2%	38.5%
Total	100.0%	100.0%	100.0%

($p<0.001$)

Students who established a peer associations with other students were more likely to be retained into the second semester. Of those who persisted into the second semester, 85.8% had one or more peer associations, and 14.2% were soloists. Among those who did not persist into the second semester 69.3% had one or more peer associations, and 30.7% were soloists. The differences in the proportion of students with peer association between the groups were statistically significant ($p<0.001$).

Significant differences surfaced in the relationship between student profiles and peer associations. Differences in peer associations are correlated with students' academic outcomes in the first semester. By examining retention into the second semester a possible correlation between student outcomes and peer associations, over time, has begun to emerge. The next section evaluates the relation between students peer association and their enrollment during the study period.

Cumulative Semesters Peer Association

The study group cohort, N=3,128 students' enrollment records were tracked at the tested college over the length of the study period (13 concurrent Semesters). Enrollment intensity for each student varied, such that over the 13 semesters the mean enrollment was equal to 4.1 semesters, with a median of 4.0 semesters. The peer associations were evaluated according to 11 categories (Cumulative number of Peer associations within the entire study cohort, Race/ethnicity, Gender, Late registration, Schedule changes, Single course enrollments, Student course load, Cumulative GPA, Cumulative credits earned, Cumulative credits attempted and Graduation). What will be seen of the study cohort over their enrollment period (cumulatively) is that the rate of peer association varied widely among individual students. On average, student exposure to peers (two or more classes) declines from the first semester (5.4%) to lower rates, ranging between 3.7% and 4.7%, over the remainder of the study period. While about two-thirds of the students (64.2%) were soloists for at least one semester during the study period, some soloists (10.1%) managed to avoid any peer association over the entire study period. The remaining 35.8% of the study cohort were never soloists.

Entire cohort peer association.

The average student in the study cohort had between 73.2 and 108.9 unique classmates each semester during the study period. Of these classmates, the average peer association was quite low, only 5.8 unique students in the first semester. Afterwards between 2.8 and 4.8 unique students were encountered, which ranged between 3.7% and 4.7% of the unique classmates. This information is presented in Table 18.

Table 18

Per Student Average Number of Unique Classmates during the study period and the Number and Percentage of Peer Associations

Semester	Students	Average Unique Classmates	Met once	2 or more classes	% Peers, 2 or more classes	3 or more classes	% Peers, 3 or more classes
1	3,128	106.8	101.0	5.8	5.4%	3.2	3.0%
2	2,383	103.4	99.6	3.8	3.7%	1.2	1.2%
3	1,848	105.7	101.8	3.8	3.6%	1.0	1.0%
4	1,532	99.4	95.8	3.6	3.6%	0.8	0.8%
5	1,163	90.7	87.1	3.5	3.9%	0.8	0.8%
6	856	110.8	106.1	4.8	4.3%	0.6	0.6%
7	577	108.9	104.1	4.8	4.4%	0.5	0.4%
8	399	106.1	101.1	5.0	4.7%	0.5	0.5%
9	263	104.5	100.2	4.4	4.2%	0.6	0.6%
10	195	83.2	79.5	3.7	4.5%	0.9	1.0%
11	161	100.6	96.6	4.0	4.0%	0.7	0.7%
12	118	86.6	83.4	3.3	3.8%	0.7	0.8%
13	92	73.2	70.4	2.8	3.8%	0.9	1.2%

After the first semester a rapid decline can be seen in the percentage of students who had three or more peer associations. The average student had 3.2 peer associations with students s/he met three or more times in the first semester. The peer association rate declines to 1.2 and 1.0 in the second and third semester, respectively. For the remainder of the study period the number peer associations is below 1.0.

Figure 3 reports the distribution of the number of peer associations a student may have in each semester.

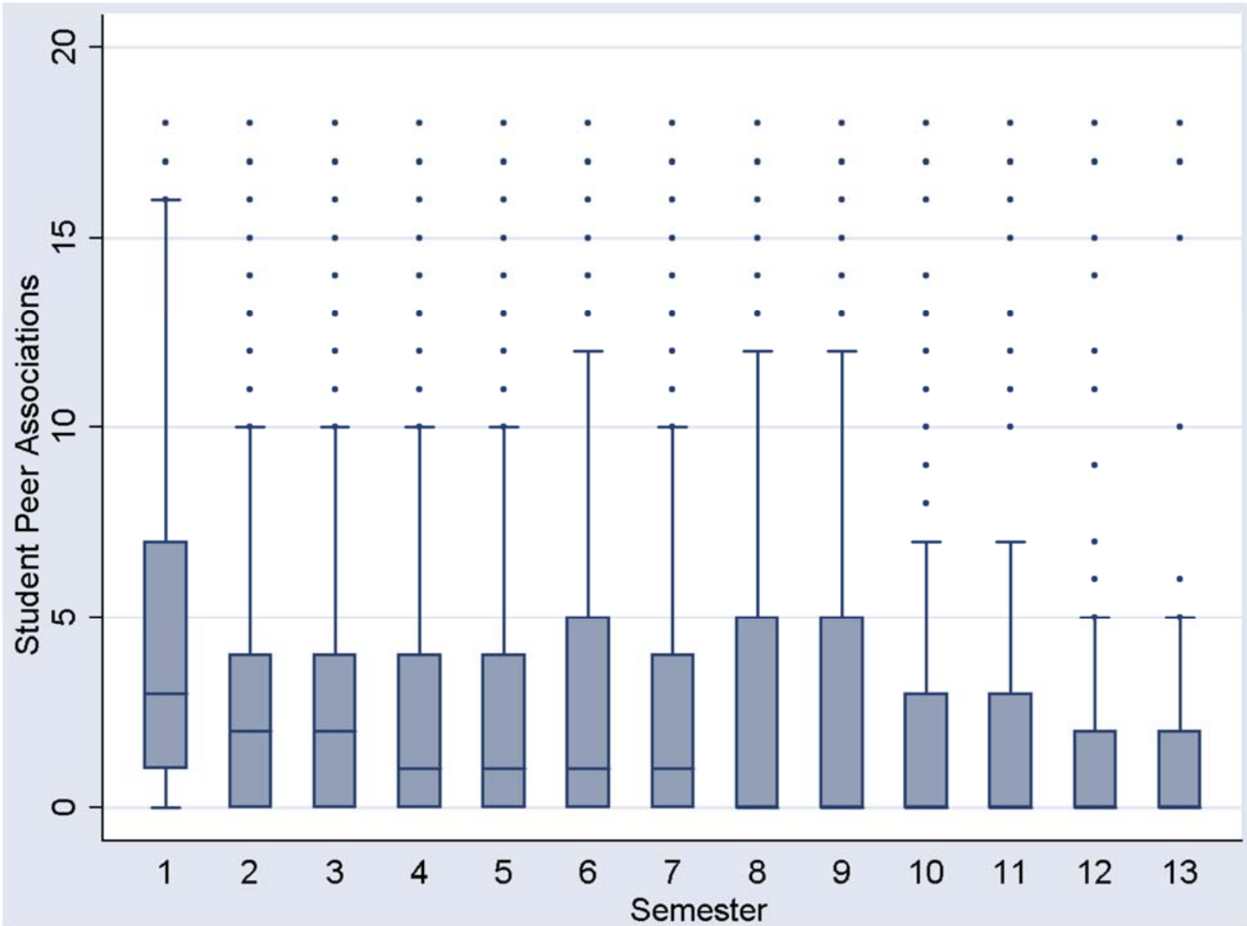

Figure 3. Distribution of Student Peer Association for Each Semester

Students are more likely to have more peer associations (between 1 and 2 more) in the first semester than in later semesters. From the second semester onwards individual students' semestrial peer associations steadily decline. Conversely, the proportion of soloists in a given semester steadily increases over time, as does the proportion of students with only one peer association. Beginning in the fourth semester, and afterwards, fewer than 50% of the students have a semestrial peer association that exceeds one peer student.

Student enrollment and the number of semesters as a soloist is a function of the number of semesters a student enrolled. Table 19 reports the number of active semesters a student enrolls and the number of semesters in which a student was a soloist.

Table 19

Number of Semesters Enrolled and Number of Semesters a Student is a Soloist throughout the Study Period

		Number of Semesters Enrolled as a Soloist									
Total Semesters Enrolled	Total Students	0	1	2	3	4	5	6	7	8	9+
1	590	405	185								
2	508	253	196	59							
3	358	139	119	78	22						
4	431	139	141	99	34	18					
5	342	88	112	70	44	16	12				
6	340	60	99	97	41	20	13	10			
7	233	25	59	62	42	28	10	3	4		
8	165	10	36	42	35	20	8	8	2	4	
9+	161	1	12	24	24	30	22	21	9	11	7
	3,128	1,120	959	531	242	132	65	42	15	15	7

Of the students in the study cohort, 64.2% (n=2,008) have been soloists for one or more semesters, and 10.1% of the students (n=316) were soloists throughout their enrollment period. Out of the study cohort, 27% (n=843) were soloists at least half of their enrollment period. However, when compared with students who earned a degree by the end of the study period (n=770) there are 6.6% (n=51) who were soloists at least for half of their enrollment period. Among students who earned a degree only 1.8% (14 out of 770) were soloists throughout their entire enrollment period, compared with those 10.1% (n=316) soloist students in the study cohort.

Student demographics and peer association.

Race/Ethnicity

Table 20 reports for race/ethnic profile the average number of active semesters a student in each category enrolled throughout the study period; as well as the proportion of average number of semesters enrolled in which a student was a soloist.

Table 20

Average Number of Semesters Enrolled, and Percentage of Semesters a Student is a Soloist throughout the Study Period, By Race/Ethnic Profile.

		Asian (n=460)	AA (n=533)	Hispanic (n=1,044)	White (n=410)	Intl (n=367)	Total (N=3,128)
Average Semesters► Enrolled		4.1	3.9	4.0	4.4	4.2	4.1
Semesters as Soloists	0	31.7%	34.8%	39.4%	33.2%	30.8%	35.8%
	1	32.9%	30.6%	29.3%	32.7%	29.2%	30.7%
	2	16.5%	18.1%	17.2%	16.5%	19.7%	17.0%
	3	9.1%	8.6%	6.4%	8.6%	9.2%	7.7%
	4	4.0%	4.1%	4.0%	3.8%	3.7%	4.2%
	5	2.7%	1.2%	1.8%	1.4%	4.4%	2.1%
	6	1.8%	1.0%	1.0%	2.2%	1.7%	1.3%
	7	0.3%	0.6%	0.3%	0.8%	0.3%	0.5%
	8	0.6%	0.6%	0.4%	0.5%	1.0%	0.5%
	9+	0.3%	0.4%	0.2%	0.3%	0.0%	0.2%
Total		100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

International students were more likely to be soloists for one or more semesters throughout the study period: 69.2% are soloists for at least one semester, compared with only 64.2% of all students in the study cohort. The average number of active semesters enrolled varied slightly by race/ethnic profile, and remained between 3.9 semesters (African American students) and 4.4 semesters (white non-Hispanic students). Hispanic students were most likely to never be soloists (39.4%), followed by African American students (34.8%), whites (33.2%0 and Asians (31.7%).

From a different perspective, Table 21 reports the distribution of peer associations per capita by race/ethnicity category throughout the student's enrollment. Per capita in this context refers to the percentage of peer associations observed of each student in each race/ethnicity group throughout the study period.

Table 21

Distribution of Peer Associations Per Capita by Race/ethnicity throughout the Study Period

	% Peer Associations Per Capita	95 % Confidence Interval
All	7.6	7.22-8.01
Asian	8.9	7.54-10.26
Hispanic	7.7	6.94-8.37
AA	5.7	5.20-6.26
Whites	8.4	7.36-9.52
Int'l	8.2	6.72-9.59

p=0.0006

When viewed in this way Asians are seen as having the highest percentage of peer associations throughout the study period (8.9%) and African Americans had the lowest peer associations per capita, 5.7%. Using Bonferroni's post hoc test Table 22 reports the significant level of the differences in peer association per capita among race/ethnic groups

Table 22

Significant Level of Differences of Peer Associations Per Capita between Race/Ethnicity Groups throughout the Study Period

	Int'l	Asian	AA	Whites	Hispanics
Asian	n/s				
AA	0.045	0.001			
Whites	n/s	n/s	0.006		
Hispanics	n/s	n/s	0.022	n/s	

Bonferroni post hoc test

African American students had significantly lower peer association per capita than all other race/ethnic groups. The African American peer association level (5.7%) was 1.9% lower than the overall cohort (7.6%), and between 2.0% and 3.2% lower than other Race/Ethnic groups.

The differences become more notable when one examines per capita differences among students by graduation status at the end of the study period. Table 23 reports the distribution of peer associations per capita by race/ethnicity category for those students who did not graduate by the end of the study period.

Table 23

Distribution of Peer Associations Per Capita by Race/ethnicity throughout the Study Period (for non-graduates)

Non-Graduates (n=2,358)	% Peer Associations Per Capita	95 % Confidence Interval
All	7.0	6.52-7.53
Asian	8.9	7.13-10.75
Hispanic	7.3	6.41-8.22
AA	5.0	4.36-5.63
Whites	7.4	5.90-8.83
Int'l	7.5	5.33-9.61

p=0.0041

Among non-graduates, Asians maintained the highest percentage of peer associations throughout their enrollment (8.9%), and Hispanics, Whites and international students maintained peer associations per capita between 7.3% and 7.5%. Similar to the first semester peer association pattern, African Americans had the lowest peer associations per capita (5.0%). Using Bonferroni's post hoc test Table 24 reports the significant level of the differences in peer association per capita among race/ethnic groups.

Table 24

Significant Level of Differences of Peer Associations Per Capita between Race/Ethnic Groups throughout their Enrollment (for non-graduates)

	Int'l	Asian	AA	Whites	Hispanics
Asian	n/s				
AA	n/s	0.002			
Whites	n/s	n/s	n/s		
Hispanics	n/s	n/s	0.036	n/s	

Bonferroni post hoc test

African American students had significantly lower peer association per capita (5.0%) compared with Hispanic and Asian students, 7.3% and 8.9% respectively.

Among students in the study cohort who graduated before the end of the study period (n=770), peer association per capita is notably different. Table 25 reports the distribution of peer associations per capita by race/ethnicity category throughout the student's enrollment, for graduates.

Table 25

Distribution of Peer Associations Per Capita by Race/ethnicity throughout their Enrollment (for graduates)

Graduates (n=770)	% Peer Associations Per Capita	95 % Confidence Interval
All	9.4	9.05-9.79
Asian	8.8	7.83-9.76
Hispanic	8.9	8.41-9.33
AA	8.4	7.70-9.05
Whites	10.9	9.71-11.99
Int'l	9.4	8.54-10.35

p=0.0009

Graduates had a higher peer association per capita compared with non-graduates, 9.4% versus 7.0%, respectively. Graduates had a higher peer association per capita rate than non-graduates across all race/ethnic groups (with the exception of Asians, who had similar rates, 8.8% and 8.9%, respectively). White graduates had the highest percentage of peer associations throughout their enrollment (10.9%). All other race/ethnic groups maintained a peer association per-capita rate of 8.4% to 9.4%. African Americans who earned a degree had a peer association per capita rate of 8.4% (compared with 5.0% for non-degree completers). Using Bonferroni's post hoc test Table 26 reports the significant level of the differences in peer association per capita among race/ethnic groups.

Table 26

Significant Level of Differences of Peer Associations Per Capita between Race/Ethnic Groups throughout the Enrollment (for graduates)

	Int'l	Asian	AA	Whites	Hispanics
Asian	n/s				
AA	n/s	n/s			
Whites	n/s	n/s	0.007		
Hispanics	n/s	n/s	n/s	0.017	

Bonferroni post hoc test

All race/ethnic groups (with the exception of Asian students) had a higher rate of peer association per capita among graduates. African Americans who completed a degree program had a similar peer association per capita rate to all ethnic groups, with the exception of white (non-Hispanic) students. Hispanic students who graduated also had a significantly lower peer association per capita compared with white (non-Hispanic students).

Gender

Table 27 reports the number of semesters a student enrolled during the study period and the number of semesters a student attended as a soloist by gender.

Table 27

Number of Semesters Enrolled In the Study Period, and Number of Semesters Enrolled As a Soloist (Female Students Only)

			Number of Semesters Enrolled as a Soloist									
		Total Females	0	1	2	3	4	5	6	7	8	9+
Total Semesters Enrolled	1	362	240	122								
	2	309	147	131	31							
	3	210	82	68	51	9						
	4	260	93	87	51	17	12					
	5	206	54	72	36	26	10	8				
	6	229	45	68	58	30	14	8	6			
	7	145	14	33	40	32	16	6	2	2		
	8	105	6	16	32	24	14	5	4	2	2	
	9+	101	1	10	15	13	22	12	11	4	7	6
		1,927	682	607	314	151	88	39	23	8	9	6

The differences in soloist rates among female and male students was negligible. (For a similar listing of male students see appendix B.) Two-thirds (64.5%) of the female students (n=1,245) were soloists for one or more semesters, compared with 63.5% for male students, and 64.2% for all students in the study cohort. Within the female population, 10.3% (n=198) were soloists throughout their enrollment, compared with 10.2% for male students, and 10.2%for the entire study cohort. Based on gender there were no statistically significant differences in the number of semesters enrolled as a soloist. Out of all female students, 26.3% (n=506) were soloists at least half of their enrollment period, compared with 27% for the entire student cohort.

Behavioral Indicators and Peer Association

On-Time vs. Late Registration

For students who were classified as Late Registrants, Table 28 compares the number of semesters for which a student was enrolled and the number of semesters enrolled as a soloist.

Table 28

Number of Semesters Enrolled as Late Registrant, and Number of Semesters Late Registrants Were Soloists, throughout the Study Period

			Number of Semesters Enrolled as a Soloist									
		Late Registrants	0	1	2	3	4	5	6	7	8	9+
Total Semesters Enrolled	1	129	84	45								
	2	104	48	38	18							
	3	64	24	19	16	5						
	4	90	12	38	29	9	2					
	5	71	9	28	18	9	4	3				
	6	89	8	27	31	10	4	5	4			
	7	62	2	11	17	14	12	2	1	3		
	8	46	2	7	11	15	6	1	2	1	1	
	9+	63	0	4	9	7	10	12	10	3	6	2
		718	189	217	149	69	38	23	17	7	7	2

Among students in the study cohort, 23.0% percent (n=718 out of 3,128 students) were late registrants once or more throughout their enrollment. Late registrants were more likely to have been a soloist 73.7% (n=589) when compared with all students (64.2%) in the study cohort.

Furthermore, of those students who were late registrants more than once (n=106), nearly all 92.5% (n=98) were soloists for one or more semesters. Students who were soloists in every semester they attended represented a slightly greater proportion among the late registrants when compared to the entire study cohort, 11.6% (n=83) and 10.1% of all students (n=316), respectively. Out of the late registrants 24.7% (n=177) were soloists at least half of their enrollment period, compared with 27% in the entire study cohort. In addition, late registrant's graduation rate was lower, 20.3% (n=146), compared with 24.6% of the entire study cohort (n=770).

Schedule Changers

For students who were classified as Schedule Changers, Table 29 compares the number of semesters for which a student was enrolled and the number of semesters enrolled as a soloist.

Table 29

Number of Semesters Enrolled and Number of Semesters as Soloist throughout the Study Period (for Schedule Changers)

		Number of Semesters Enrolled as a Soloist									
Total Semesters Enrolled	Schedule Changers	0	1	2	3	4	5	6	7	8	9+
1	258	188	70								
2	331	166	130	35							
3	279	102	96	65	16						
4	340	104	112	81	28	15					
5	288	74	97	58	34	16	9				
6	284	53	80	82	35	17	9	8			
7	213	24	55	56	37	25	9	3	4		
8	149	10	32	41	34	18	4	7	1	2	
9+	157	1	12	24	22	29	22	21	9	10	7
	2,299	722	684	442	206	120	53	39	14	12	7

Among students in the study cohort, 73.5% percent (n=2,299 out of 3,128 students) performed a schedule change on or after the first day of class at least once during their enrollment. Schedule changers were more likely to have been a soloist 68.6% (n=1,577) when compared with all students (64.2%) in the study cohort. Of those students who changed their schedule more than once (n=1,302), only 78.7% (n=1,024) were soloists for one or more semesters. Students who were soloists in every semester they attended represented a slightly smaller proportion among the schedule changers when compared to the entire study cohort, 7.2% (n=166) and 10.1% (n=316), respectively. Among schedule changers 21.6% (n=497 out of 2,299) were soloists at least half of their enrollment period, compared with 27% of the entire study cohort. Schedule changers showed a difference in their graduation rate, 26.9% (n=618), compared with 24.6% of the entire study cohort (n=770).

Student Academics and Peer Association.

Single Course Takers

The proportion of soloists, and the number of semesters in which a student is a soloist, is impacted by the number of classes enrolled in a given semester. Students who take only one course in a particular semester, by default, are soloists in that semester. Table 30 reports the number and percentage of single course enrollees during the study period.

Table 30

Number (and Percentage) of Single Course Enrollees, by Semester

Semester	Enrollment	Single Course Enrollees	Percent
1	3,128	161	5.1%
2	2,383	228	9.6%
3	1,848	193	10.4%
4	1,532	194	12.7%
5	1,163	211	18.1%
6	856	184	21.5%
7	577	150	26.0%
8	399	136	34.1%
9	263	81	30.8%
10	195	70	35.9%
11	161	55	34.2%
12	118	48	40.7%
13	92	26	28.3%
Total	12,715	1,737	13.7%

In the first semester 5.1%of students took only one course and were therefore soloists. Over time, students have a tendency to reduce their course load. From the second semester onwards, the proportion of students taking only one course in a particular semester is increasing from 9.6 percent up to 40.7 percent (in the 12th semester). Between the third and the eighth semester, single course takers (soloists) represent between one tenth (1/10th) and one third (1/3rd) of students. While the students did not develop new peer associations when they enroll in only one class, it is possible that they retained peer associations from a previous semester (or semesters). Table 31 examines the rate at which peer networks are maintained for single course takers.

Table 31

Number (and Percentage) of Students who take a Single course and Maintain Peer-Associations from Prior Semesters

Semester	Single Course Enrollees	# w/ Peer(s) from Previous Semesters	% w/ Peer(s) from Previous Semesters
1	161	x	x
2	228	88	38.6%
3	193	59	30.6%
4	194	59	30.4%
5	211	62	29.4%
6	184	67	36.4%
7	150	54	36.0%
8	136	45	33.1%
9	81	23	28.4%
10	70	18	25.7%
11	55	16	29.1%
12	48	19	39.6%
13	26	8	30.8%
Total	1,737	518	29.8%

Less than one-third (29.8%) of single course takers maintained peer associations from prior semesters. Less than two-fifths (38.6%) of the 228 students enrolled in a single course during their second semester maintained peer associations. Thereafter, we can see a downward trend among single course takers in the number of peer associations they maintained from their previous enrollment. Overall, only 29.8% of the students who took a single course are seen to have peer associations from a prior semester. The implication is that even if the student did develop a cumulative peer network created throughout their enrollment, being a soloist for a given semester weakens their network and reduces their ability to maintain peer associations.

Credits attempted and credits earned

One of this study's hypotheses is concerned with if student peer associations has an effect on the academic outcome of number of cumulative credits earned. Table 32 reports the students' cumulative credits earned by the number of semesters in which a student has been a soloist.

Table 32

Students' Average Credits Earned, by Number of Semesters Enrolled and Number of Semesters as a Soloist

		Number of Semesters Enrolled as a Soloist									
Total Semesters Enrolled	Avg Credits Earned	0	1	2	3	4	5	6	7	8	9+
1	2.72	3.08	1.94								
2	8.69	10.57	7.31	5.19							
3	16.14	19.53	16.20	10.53	14.27						
4	31.35	45.50	29.58	21.65	15.15	20.00					
5	42.46	51.53	49.09	36.46	29.32	16.81	31.33				
6	51.57	60.50	58.00	52.29	42.34	33.20	20.85	†			
7	54.55	67.00	62.17	55.97	50.36	41.96	†	†	†		
8	58.35	†	67.39	62.98	57.94	52.05	†	†	†	†	
9+	61.34	†	81.50	66.88	68.71	63.13	59.32	44.43	†	49.45	†

† n≤10

Among students with equal number of semesters enrolled, more credits are earned the less often a student is a soloist. For example, within Table 32, among students who enrolled for a total of six semesters, students who were never soloists earned an average of 60.50 credits, and students who were soloists longer, (i.e.,5 semesters) earned fewer credits (20.85) on average. This is important because nearly two-thirds (64.2%) of the study cohort (n=2,008) were soloists for one or more semesters, and 15.7% of those (n=316 out of 2,008) were soloists throughout their entire enrollment.

Similarly, Table 33 reports the students' cumulative credits attempted by the number of semesters in which a student has been a soloist.

Table 33

Students' Cumulative Number of Credits Attempted, by Number of Semesters Enrolled and Number of Semesters a Student is a Soloist

		Number of Semesters Enrolled as a Soloist									
Total Semesters Enrolled	Avg Credits Attempted	0	1	2	3	4	5	6	7	8	9+
1	6.46	7.42	4.35								
2	14.73	18.28	12.28	7.63							
3	25.04	30.92	25.23	16.15	18.32						
4	40.64	52.76	40.77	32.00	24.91	23.28					
5	52.15	61.67	58.25	48.31	36.8	28.94	35.08				
6	61.64	70.42	68.57	62.31	53.68	41.50	30.77	†			
7	67.41	79.88	75.05	69.18	64.48	51.32	†	†	†		
8	73.13	†	82.42	77.98	73.09	66.20	†	†	†	†	
9+	81.36	†	101.75	88.17	86.33	84.50	81.00	69.00	†	62.00	†

Note: † n≤10

Similar to average credits earned, there is an inverse relationship between the number of semesters a student is a soloist and the number of credits attempted. For example, students who enrolled for six semesters, and were never soloists, attempted an average of 70.42 credits, while students who were soloists for five out of six semesters attempted only 30.77 credits, on average.

The Correlation between Student Peer Association and Student Outcomes

The richness of information contained in student records systems enables a detailed analysis of peer associations, and student-level outcomes (i.e. GPA, credits attempted, credits earned), this section will explore the relationship among these and other factors such as: Behaviors, demographics, academics, and financial indicators. Further analysis will explore the impact of these factors on student-level outcomes in a specific semester, and overtime. The rate of peer-association among students is derived by identifying coincidental course enrollment each semester, and then determining to what extent a pattern develops between individuals over time. The descriptive statistics demonstrated that student peer-association varies in terms of the number of peer associations, and the intensity of peer associations. Students may encounter peers just once, or multiple times. The growth model analysis takes both into account: The analyses of student-level outcomes employs variables that evaluate the impact of peer associations, considering three levels of intensity for peer associations: two encounters, three encounters, and four or more encounters, between a student in the study cohort and a unique peer.

Semester's GPA

One of this study's hypothesis ($H2_a$) asserts that peer association is positively associated with student's semestrial GPA. Accounting both for the number and intensity of peer associations Table 34 compiles the correlation among the five different models tested in order to test this hypothesis. What will be seen is that peer-association has a consistent, and positive correlation with student semestrial GPA. The correlation remains positive and significant once additional covariates are taken into account. Groups of covariates are added in the following order: Peer association (model 1), Additional behavioral covariates (model 2), Academic covariates (model 3), Demographic covariates (model 4), and Financial covariates (model 5).

Table 34

Students' Semestrial GPA by Peer Association, Additional Behaviors, Academic, Demographic, and Financial Indicators (Models 1 through 5)

Y=Semestrial GPA	Model 1	Model 2	Model 3	Model 4	Model 5
Behavioral Indicators					
Peer Associations (Encountered Twice)	0.033 *** (0.004)	0.032 *** (0.004)	0.013 *** (0.004)	0.012 ** (0.004)	0.012 ** (0.004)
Peer Associations (Encountered 3 times)	0.032 *** (0.006)	0.033 *** (0.006)	0.022 *** (0.006)	0.017 ** (0.006)	0.016 * (0.006)
Peer Associations (Encountered 4+ times)	0.044 *** (0.005)	0.042 *** (0.005)	0.023 *** (0.005)	0.019 *** (0.005)	0.019 *** (0.005)
Late Registration		-0.216 *** (0.042)	-0.198 *** (0.042)	-0.177 *** (0.042)	-0.167 *** (0.042)
Change Schedule		-0.176 *** (0.02)	-0.202 *** (0.021)	-0.203 *** (0.02)	-0.205 *** (0.02)
Return from Stopout		0.145 *** (0.044)	0.250 *** (0.046)	0.243 *** (0.046)	0.248 *** (0.046)
Academic Indicators					
Need Writing (1st day of semester)			-0.134 *** (0.033)	-0.104 *** (0.033)	-0.097 ** (0.032)
Need Math (1st day of semester)			-0.248 *** (0.036)	-0.192 *** (0.035)	-0.182 *** (0.035)
Total Credits Earned (1st day of semester)			0.010 *** (0.001)	0.009 *** (0.001)	0.009 *** (0.001)
Credits attempted in the semester			-0.024 *** (0.004)	-0.020 *** (0.004)	-0.021 *** (0.004)
Number of Classes in the semester			0.108 *** (0.012)	0.119 *** (0.012)	0.119 *** (0.012)
Part Time in the semester			-0.030 (0.034)	-0.028 (0.034)	-0.054 (0.034)
Demographic Indicators					
Age				0.028 *** (0.002)	0.028 *** (0.002)
Male				-0.194 *** (0.035)	-0.196 *** (0.034)
Race/Ethnic Profile (Reference=White)					
Int. Student				0.415 *** (0.053)	0.330 *** (0.057)
Asian				-0.050 (0.064)	-0.037 (0.063)
Hispanic				-0.397 *** (0.055)	-0.362 *** (0.055)
African Americans				-0.514 *** (0.062)	-0.479 *** (0.061)
Other				-0.330 *** (0.059)	-0.312 *** (0.058)
Financial Indicators					
Paid with Aid (Any Portion of Tuition)					0.430 *** (0.048)
Paid with Self (Any Portion of Tuition)					-0.031 (0.031)
Percent Tuition Paid using Aid					-0.434 *** (0.062)
Pell or Tap recipient					-0.170 *** (0.041)
Time (Semesters)	-0.029 *** (0.005)	-0.035 *** (0.005)	-0.088 *** (0.008)	-0.087 *** (0.008)	-0.082 *** (0.008)
Constant	2.163 *** (0.025)	2.266 *** (0.027)	2.225 *** (0.055)	1.776 *** (0.095)	1.866 *** (0.102)

Note: Coefficients (and Standard Errors) reported, *p≤0.05, **p≤0.01, ***p≤0.001

These results report the estimated change in semester GPA in any given semester attributable to peer association intensity. Model 1 shows that student peer association is correlated with an increase in the student's semestrial GPA. Having a single peer association (encountered classmate 2 times) is associated with to an increase of 0.033 GPA points in a student's semestrial GPA when compared with the GPA of a soloist student ($p \leq 0.001$). Similarly, the correlation of student peer association at a higher exposure (encountered classmate three times) is also positive, but changes little from that of peers who were encountered only twice (0.032 vs. 0.033). However, at higher levels of peer intensity (encountered classmate four or more times) the association is strongest with student's semestrial GPA, and it is correlated with an increase of 0.044 GPA points for each high intensity peer association ($p \leq 0.001$). The correlation of high intensity peer associations is also stronger, as reflected in the standardized score ($Z=9.210$) versus $Z=9.060$ and $Z=5.180$ for peers encountered only two or only three times, respectively.

The influence of peer association may be partially attributed to additional behavioral covariates. Hypothesis $H2_a$ asserts that other student-level behavioral indicators may be correlated student's GPA. Model 2 evaluates the influence of Student's Peer Association and additional behavioral covariates (Late registration, Schedule changes and Return from a stopout) on their semestrial GPA. The influence of peer association intensity in Model 2 is consistent (positive, $p \leq 0.001$) even after controlling for additional behavioral covariates: Two encounters =0.032, three encounters =0.033 and four or more encounters =0.042 GPA points.

Late registration in a given semester is associated with a 0.216 points decline in the student semestrial GPA ($p \leq 0.001$). Class schedule changes, on or after the first day of classes, is associated with a 0.176 points decline in semestrial GPA. Standardized scores reveal that the relative correlation of late registration and schedule changes are similar to that of student peer-associations with Z=5.140 and Z=8.620, respectively. The standardized score for student peer association varies between 5.280 and 8.720. Students who returned from a stopout are likely to have a 0.145 points increase in semestrial GPA in the semester they returned. ($p \leq 0.001$).

Model 3 evaluated the influence of student peer association while controlling for additional behavioral and academic covariates. In the presence of academic covariates the correlation of student peer association with student semestrial GPA was consistent (positive, $p \leq 0.001$), with gains in semestrial GPA of 0.013, 0.022 and 0.023 GPA points for each peer association, two, three, and four or more encounters, respectively. The standardized scores though are lower, Z=3.210, Z=3.350 and Z=4.290, respectively (without academic indicators, the standardized scores were Z=8.830, 5.280 and 8.720, respectively). These changes in slope are reasonable - the number of peers associations and peer encounter intensity is related to the number of classes (and credits) a student attempts. Part time students, and students who attempt fewer classes (or credits) would have fewer opportunities to establish peer associations[2]. The slope of those academic covariates would only partially represent the influence previously recorded based on peer associations alone. There were similar declines in in the expected semestrial GPA related to late registration (-0.198 points) and schedule changes (-0.202 points, $p \leq 0.001$). A return from a stopout is associated with an increase of 0.250 points in a student's GPA for the semester in which the student returned. Student remedial education needs was negatively associated with semestrial GPA, with a decline of 0.134 GPA points for remedial Writing is and 0.248 points for remedial Math. The total number of credits earned by a student

[2]Table 15 provides a detailed overview of the number of classes taken by a student and the number of peer associations in the first semester. Figure 2 outlines the relationship between the number of classes taken and the percentage of students who are soloists using two different definitions (two and three peer encounters)

prior to first day of classes was associated with an increase of 0.010 GPA points (e.g. each academic course (typically 3 credits) which would result in an increase of 0.030 points to a student's semestrial GPA).

The effect of course load is expressed using two variables: classes attempted and credits attempted. Student GPA increased by 0.108 points for each additional class attempted ($p \leq 0.001$). A slight decline in GPA (0.024 points, $p \leq 0.001$) can be attributed to each credit attempted. For a typical credit bearing course (3 credits), the net influence would be positive, adding an additional 0.036 points to the semestrial GPA (0.108-0.024*3=0.036 points). Being a part time student was not significantly related (although negative) with student expected GPA.

Model 4 reports that peer association relationship to students' GPA is significant in the presence of additional behavioral and academic covariates. However, the effect of peer association may be partially attributed to demographic covariates. Male students, minorities and older students earned lower GPA compared with White (non-Hispanic), female, and traditional age students. Model 4 shows that when student-level demographic indicators are accounted for, the correlation of student peer association remains in place. Peer association is correlated with increases students' expected semestrial GPA by 0.012, 0.017 and 0.019 points for each additional peer association (two, three, and four or more encounters, respectively ($p \leq 0.006$) compared with soloist students. Additional behavioral covariates reflect similar effects on student semestrial GPA: Late registration and schedule changes are associated with expected decreases in semestrial GPA (-0.177 and -0.203 points, respectively ($p \leq 0.001$)). Returning after a stopout is positively associated with semestrial GPA (increase of 0.243 points, $p \leq 0.001$). Students' demographics are associated with changes in the expected semester GPA: age is associated with a small increase in the expected GPA (0.028 points per year), and male students are expected to have -0.194 points decrease in GPA compared with female students ($p \leq 0.001$). International students (student visa) are expected to earn a GPA that is 0.415 points higher than other students (when all other covariates held constant). Minority students (with the exception of Asians) are expected to have lower semestrial GPA: Hispanic, African American and other minority student groups' expected GPA is 0.397, 0.514 and 0.330 points lower than white, non-Hispanic students ($p \leq 0.001$). The negative influence for African American students is stronger than other minority groups, and seems to correspond with their lower rates of peer-association[3]. Overall, the influence of the number of semesters enrolled remains similar: For every semester a student enrolls, there is a slight decline, in the semestrial GPA (-0.087 points, $p \leq 0.001$). The influence of academic indicators on student semestrial GPA remains consistent with previous findings. But, the influence of student's part-time status (which remains insignificant, -0.028, $p=0.427$) may reflect on other student's qualities. Specifically, being part-time may impact a student's eligibility for various financial aid programs (federal, state or regional, or institutional aid).

Model 5 adds reports that the correlation of student peer association remains consistent even when financial indicators are considered. Peer association increases expected student GPA by 0.012, 0.016 and 0.019 points for each peer association a student has (two, three, and four or more encounters, respectively) compared with soloist students. The association with other behavioral indicators remains similar: Late registration and schedule changes are negatively associated with expected decline in semestrial GPA (-0.167 and -0.205 points, respectively ($p \leq 0.001$)), and returning from a stopout is positively associated with increased semestrial GPA (0.248 points, $p \leq 0.001$). The use of financial aid to pay for a student's tuition was associated

[3] Peer association per capita is reported under table 6 (first semester) and table 21 (cumulatively)

with increased semestrial GPA. The influence was related to an expected increase of 0.430 semestrial GPA points ($p \leq 0.001$). This increase was important across the study group in that most students received some form of aid at least once during the study period. Beginning with the first semester 71.9% (n=2,251) of the study group used financial aid to pay for some or all of their first semester's tuition. During the study period nearly two thirds (64.5%) of the semestrial tuition accounts for the study cohort was paid partially, or completely, using financial aid. The covariate "percent tuition paid using aid" further estimates relationship between financial aid and GPA. When tuition was completely paid for with financial aid the net effect was to neutralize the GPA benefits of financial aid. (-0.434 for having 100% of the tuition paid using aid versus 0.430 semestrial GPA points for paying less than full tuition with financial aid). During the study period, an average of 58.8% of student's tuition dollars were paid using financial aid. The result was a net increase of 0.175 points in student semestrial GPA for aid recipients. Pell or TAP[4] recipient status for a given semester is associated with a decline in student GPA (-0.170 points, $p \leq 0.001$). Student eligibility for either PELL or TAP is consistent with low socio-economic status, a factor associated with several risks factors related to early departure (Bean & Metzner, 1985; Pascarella & Terenzini, 2005). Tuition paid without financial aid had an insignificant influence on student semestrial GPA. The overall influence of student academics and demographics on GPA remains similar, though student visa status had a weaker effect than previously measured resulting in an increased GPA of 0.330 points (a standardized score of 5.83), compared with an increase of 0.415 points (a standardized score of 7.79) in the previous model. It is possible that part of the influence of international student status is now captured in student financial indicators: international students are ineligible for most need-based aid

[4] A tuition only award for undergraduate students in New York State. Eligibility is based on US citizenship/permanent resident aliens, NYS residents who are attending a NYS approved college/institution full time while making academic progress towards a degree.

programs, and derive financial support mainly from non-aid resources or direct institutional assistance.

Overall, the findings also show that the association between behavioral indicators and student's semestrial GPA is influential (negative and positive effects), and remains significant in presence of other covariates. The association became weaker (but remained significant) once academic covariates were included.

Semestrial Credits Earned

Hypothesis $H2_b$ asserts that peer association is positively associated with student's Credits Earned in a given semester. Table 35 evaluates this hypothesis, accounting for the number of peer associations, and the intensity of those associations. What will be seen is that the number of peer-associations has a positive correlation with student credits earned. The correlation remains positive and significant once additional covariates are taken into account. Groups of covariates are added to the model in the following order: Peer association (Model 1), Additional behavioral (Model 2), Academic (Model 3), Demographic (Model 4) and Financial covariates (Model 5).

Table 35

Students' Credits Earned by Peer Association, Additional Behaviors, Academic, Demographic, and Financial Indicators (Models 1 through 5)

Y=Semestrial Credits Earned	Model 1	Model 2	Model 3	Model 4	Model 5
Behavioral Indicators					
Peer Associations (Encountered Twice)	0.673 *** (0.016)	0.647 *** (0.016)	0.079 *** (0.011)	0.077 *** (0.011)	0.077 *** (0.011)
Peer Associations (Encountered 3 times)	0.335 *** (0.028)	0.329 *** (0.028)	0.111 *** (0.018)	0.103 *** (0.018)	0.102 *** (0.018)
Peer Associations (Encountered 4+ times)	0.212 *** (0.021)	0.208 *** (0.021)	0.051 *** (0.015)	0.043 ** (0.015)	0.044 ** (0.015)
Late Registration		-2.319 *** (0.169)	-0.412 *** (0.106)	-0.351 *** (0.106)	-0.319 ** (0.107)
Change Schedule		-0.561 *** (0.087)	-1.316 *** (0.055)	-1.307 *** (0.055)	-1.301 *** (0.055)
Return from Stopout		-1.571 *** (0.176)	0.147 (0.111)	0.182 (0.111)	0.208 † (0.111)
Academic Indicators					
Need Writing (1st day of semester)			-0.293 *** (0.084)	-0.251 ** (0.083)	-0.248 ** (0.083)
Need Math (1st day of semester)			-0.215 * (0.088)	-0.103 (0.087)	-0.093 (0.087)
Total Credits Earned (1st day of semester)			0.019 *** (0.002)	0.019 *** (0.002)	0.021 *** (0.002)
Credits attempted in the semester			0.746 *** (0.009)	0.749 *** (0.009)	0.745 *** (0.009)
Number of Classes in the semester			0.143 *** (0.033)	0.162 *** (0.033)	0.153 *** (0.033)
Part Time in the semester			-0.069 (0.091)	-0.070 (0.091)	-0.114 (0.092)
Demographic Indicators					
Age				0.045 *** (0.006)	0.044 *** (0.006)
Male				-0.315 *** (0.083)	-0.310 *** (0.083)
Race/Ethnic Profile (Reference=White)					
Int. Student				0.687 *** (0.126)	0.568 *** (0.135)
Asian				0.112 (0.153)	0.125 (0.151)
Hispanic				-0.499 *** (0.132)	-0.468 *** (0.132)
African Americans				-0.646 *** (0.148)	-0.612 *** (0.148)
Other				-0.531 *** (0.141)	-0.521 *** (0.14)
Financial Indicators					
Paid with Aid (Any Portion of Tuition)					0.643 *** (0.114)
Paid with Self (Any Portion of Tuition)					-0.145 † (0.085)
Percent Tuition Paid using Aid					-0.611 *** (0.151)
Pell or Tap recipient					-0.263 ** (0.098)
Time (Semesters)	0.010 (0.018)	0.049 ** (0.018)	-0.122 *** (0.017)	-0.129 *** (0.017)	-0.132 *** (0.017)
Constant	4.897 *** (0.095)	5.361 *** (0.102)	-0.172 (0.141)	-0.995 *** (0.231)	-0.723 ** (0.253)

Note: Coefficients (and standard errors) reported, †p<0.1, *p≤0.05, **p≤0.01, ***p≤0.001

In Model 1, student peer association is correlated with an increase in the student's credits earned. The influence of a single peer association (encountered 2 times) is equivalent to an increase of 0.673 credits earned ($p \leq 0.001$) in a given semester compared with that of a soloist student. Similarly, the correlation of student peer association at higher exposures was 0.335 credits increase ($p \leq 0.001$) when a unique peer association was encountered three times, and 0.212 credits increase ($p \leq 0.001$) when there were four or more encounters. Surprisingly, the correlation of peer association on credits earned seems to decline with increased peer association intensity.

Further evaluation (Models 2 and 3) shows that increases in semester credits earned are partially attributed to other covariates. The influence of peer association in Model 2 remains similar once the additional behavioral covariates of Late registration, Schedule changes, and Return from stopout are added. All were associated with fewer semester credits earned. Late registrants earned 2.319 fewer credits ($p \leq 0.001$) than on time registrants, and schedule changers earned 0.561 fewer credits ($p \leq 0.001$) compared with students retaining their initial schedule. Students returning from stopout earned 1.571 fewer credits ($p \leq 0.001$) in the semester they returned, compared with a student who maintained continuous enrollment in the same semester.

Once academic indicators are added (model 3), the correlation of peer associations remains significant but it becomes smaller than previously estimated. Peer association increases students' credits earned (by 0.079, 0.111 and 0.051 for each peer association, two, three, and four or more encounters, respectively,) compared with soloist students (sig.≤ 0.001). Late registration and schedule changes are negatively associated with credits earned for a given semester, 0.412 and 1.316 fewer credits, respectively.

The influence of student academic covariates in any given semester is notable: Total credits earned to date, Credits attempted, and Classes attempted are all positively associated with number of credits earned. The total number of credits earned prior to first day of classes increases the number of credits earned in the semester by 0.019 per credit earned. Number of classes attempted is associated with an increase of semester credits earned of 0.143 credits ($p \leq 0.001$). The number of credits attempted in a semester increases the expected credits earned by 0.746 credits for each additional credit attempted, or an average of 2.238 additional credits earned for a typical 3 credit course. Compared with college-level students needing remedial Writing or Math, are both negatively associated with number of credits earned, -0.293 credits earned, and -0.215 credits earned, respectively.

Model 4 introduces student demographic covariates which do not change the association of behavioral or academic covariates previously measured in Model 2 and Model 3. First, the correlation of peer association to semester credits earned remains consistent. Peer association intensity increases students' credits earned by 0.077, 0.130 and 0.043 for two, three, and four or more encounters, respectively. Second, late registration and schedule changes are negatively associated with semester credits earned, 0.351 and 1.307 fewer credits, respectively ($p \leq 0.001$), while return from stopout has an insignificant effect on semestrial credits earned. In terms of student demographics, age is associated with an increase 0.045 semester credits earned per year. International students (student visa) earned 0.687 more semester credits when compared with resident students. Hispanics, African Americans and other minority students (excluding Asians) earned fewer semester credits, 0.499, 0.646 and 0.531 fewer credits, respectively ($p \leq 0.001$). Males earned 0.315 fewer semester credits than females ($p \leq 0.001$).

Lastly, with the addition of financial covariates, Model 5 shows that the correlation of student peer association with credits earned in the presence of student behavioral covariates is consistent across all models tested. Peer association intensity is associated with increases in semester credits earned of 0.077, 0.102 and 0.044 for two, three, and four or more encounters, respectively ($p \leq 0.003$). Late registrants and schedule changers earned 0.319 and 1.301 fewer semester credits, respectively. Students returning from a stopout showed an increase of 0.208 credits compared with those students maintaining continuous enrollment ($p=0.061$).

In this set of models financial covariates played a significant role. Paying partial tuition using financial aid increases credits earned by 0.643 semester credits. However, similar to the findings in previous models, when tuition was completely subsidized with financial aid the net effect becomes smaller (0.643-0.611=0.032), all other covariates held constant. Pell or TAP recipient status within a given semester is associated with 0.263 fewer credits earned ($p=0.007$); and tuition paid without any financial aid is associated with 0.145 fewer semestrial credits earned, though its influence is borderline significant ($p=0.090$).

Overall, these Models 1-5 show that student behavioral indicators are correlated with semester credits earned. The inclusion of academic level covariates changes the slopes of student behavioral indicators, yet their association with semestrial credits earned remained strong. It is possible that the number of peer associations is more crucial in the influence on student's semestrial credits earned than the intensity of the association. Late registration and schedule changes were negatively correlated with student semestrial credits earned, while return from a stopout was associated with an increase in semestrial credits earned during the semester in which the student returned.

Student Cumulative GPA

Next, hypothesis $H2_c$ asserts that peer association is positively associated with student's cumulative GPA is examined. Hypothesis $H2_c$ examines if the correlation of student peer association is cumulative, and whether it remains over time (even over those semesters a student does not have direct exposure to peers). Table 36 evaluates this hypothesis by accounting for the number of peer associations, and the intensity of those associations throughout the entire study period. What will be seen is that cumulative peer association is positively associated with cumulative GPA in ways that vary from the previous section, which examined individual semester peer-association. Variation in the influence of peer association (specifically cumulative peer association) on cumulative GPA is revealed with the introduction of additional covariates. The covariates are added to Table 36 in the following order: Peer association (Model 1), Additional behavioral (Model 2), Academic (Model 3), Demographic (Model 4) and Financial covariates (Model 5).

Table 36

Students' Cumulative GPA by Peer Association, Additional Behaviors, Academic, Demographic, and Financial Indicators (Models 1 through 5)

Y=Cumulative GPA	Model 1	Model 2	Model 3	Model 4	Model 5
Behavioral Indicators					
Cumulative Peer Associations (Encountered 2 times)	-0.001 (0.001)	0.000 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Cumulative Peer Associations (Encountered 3 times)	-0.002 (0.002)	0.000 (0.002)	0.002 (0.003)	0.002 (0.003)	0.001 (0.003)
Cumulative Peer Associations (Encountered 4+ times)	0.003 (0.002)	0.005 * (0.002)	0.007 *** (0.002)	0.006 ** (0.002)	0.006 ** (0.002)
Cumulative Late Registration		-0.082 ** (0.03)	-0.069 * (0.03)	-0.057 † (0.03)	-0.059 * (0.03)
Change Schedule		-0.036 *** (0.011)	-0.032 ** (0.011)	-0.034 ** (0.011)	-0.036 *** (0.011)
Cumulative Return from Stopout		0.227 *** (0.037)	0.194 *** (0.045)	0.207 *** (0.045)	0.236 *** (0.046)
Academic Indicators					
Ever Need Writing (1st day of semester)			-0.178 *** (0.047)	-0.151 *** (0.045)	-0.150 *** (0.046)
Ever Need Math (1st day of semester)			-0.400 *** (0.048)	-0.319 *** (0.047)	-0.314 *** (0.047)
Total Credits to date			0.000 (0.001)	0.000 (0.001)	0.000 (0.001)
Cumulative Credits attempted			-0.006 *** (0.002)	-0.005 ** (0.002)	-0.006 *** (0.002)
Cumulative Number of Classes			0.010 † (0.005)	0.010 † (0.005)	0.011 * (0.006)
Cumulative Semesters Enrolled as part time			-0.012 (0.015)	-0.012 (0.014)	-0.031 * (0.015)
Demographic Indicators					
Age				0.023 *** (0.003)	0.023 *** (0.003)
Male				-0.252 *** (0.044)	-0.256 *** (0.044)
Race/Ethnic Profile (Reference=White)					
Int. Student				0.255 *** (0.05)	0.244 *** (0.05)
Asian				0.029 (0.082)	0.027 (0.082)
Hispanic				-0.343 *** (0.07)	-0.327 *** (0.07)
African Americans				-0.499 *** (0.078)	-0.483 *** (0.078)
Other				-0.325 *** (0.074)	-0.318 *** (0.074)
Financial Indicators					
Cum. Semesters Paid with Aid (Any Portion of Tuition)					0.001 (0.007)
Cum. Semesters Paid with Self (Any Portion of Tuition)					0.049 *** (0.012)
Cumulative Percent Tuition Paid using Aid					-0.018 (0.041)
Cumulative Pell or Tap recipient					-0.005 (0.011)
Time (Semesters)	-0.011 * (0.006)	-0.013 † (0.008)	0.006 (0.018)	-0.010 ** (0.018)	-0.029 (0.019)
Constant	2.341 *** (0.023)	2.360 *** (0.024)	2.561 *** (0.035)	2.305 *** (0.098)	2.315 *** (0.100)

Note: Coefficients (and standard errors) reported, †p<0.1, *p≤0.05, **p≤0.01, ***p≤0.001

Model 1 shows no evidence that student peer association had a correlation with cumulative GPA. Neither the number nor intensity peers association effected students' cumulative GPA, unlike other models examined in this study up to this point, where peer association was correlated to academic outcomes. It is possible therefore that peers encountered in previous semesters (but are not part of a student's current enrollment pattern) have little or no correlation with the students' cumulative GPA. The next model will examine the influence of additional covariates (behavioral and academic) on student peer association.

Model 2 shows that students who encountered the same peers four or more times experienced a small but positive association with cumulative GPA (0.005 for each peer association, z=2.490, p=0.013). Cumulative late registration is associated with a decline in cumulative GPA of 0.082 for each late registration incident. Similarly, a student who changed their class schedule on or after the first day of classes is expected to have a GPA that is 0.036 lower for each semester a student initiated schedule changes. A student who returned from a stopout is likely to earn about 0.227 points higher GPA for each return from stopout. This finding is somewhat surprising, but it is possible that return from stopout corresponds with improve organizational skills and other preparations that allow the student to concentrate on their studies.

In Model 3 (once academic covariates are added to the model), the correlation between peer association and cumulative GPA is limited only to high intensity peer association. The is negative association of student late registration and schedule changes remains similar to the ones measured without academic indicators: each late registration event is associated with a decline of 0.069 points in cumulative GPA; each semester of schedule changes is associated with a decline of 0.032 points in cumulative GPA; and a return from a stopout is associated with an increase of

0.194 in a student's cumulative GPA. Meanwhile, a significant correlation with student's cumulative GPA is associated the cumulative number of classes attempted indicating an increase of 0.010 in cumulative GPA for each class attempted. This association is understandable given that peers association and peer association intensity is related to the number of classes a student attempts. Cumulative part time status is associated with a decline of 0.012 GPA points for each semester enrolled as part time student. Students who needed remedial Writing or remedial Math earned a cumulative GPA that was 0.178 and 0.400 lower, respectively, than students who entered at college level ($p \leq 0.001$).

The addition of demographic indicators in Model 4 did not change the correlation between student peer associations and student behavioral indicators. The correlation of high intensity peer associations (encountered four or more times), and a student returning from stopout on students' cumulative GPA remained positive; The correlation of late registration and schedule changes on students' GPA remained negative. Students' demographics are associated with the following changes in cumulative GPA: Age is associated with a small increase in the expected GPA (0.023 per year, $p \leq 0.001$), and male students are expected to have cumulative GPAs 0.252 lower as compared with female students ($p \leq 0.001$). This relationship is similar to that for semestrial GPA. International students (student visa) are expected to earn cumulative GPAs 0.255 points higher than residents, all other covariates held constant. Also similar to semestrial GPA, minority students (with the exception of Asians) are expected to have lower cumulative GPA: Hispanic, African American and other minority student groups' expected cumulative GPA is 0.343, 0.499 and 0.325 points lower than white, non-Hispanic students ($p \leq 0.001$). The influence of the number of semesters enrolled on student cumulative GPA is insignificant.

Once financial indicators are considered in Model 5, the correlation between students' high intensity peer association (encountered four or more times), late registration, schedule changes, returning from a stopout and students' cumulative GPA remained consistent with previous findings. The association between academic covariates and cumulative GPA remains similar to all but part time status: Cumulative part time enrollment status is associated with a decline of 0.031 points in students' cumulative GPA for each semester ($Z=-2.080$, $p=0.038$) compared with full time enrollees. While the influence of student's part time status was insignificant in Model 3 and Model 4, it may reflect on other student's qualities. Specifically, part time status influences a student's eligibility for several federal, state or regional aid programs, as well as for institutional aid. The influence of financial indicators is as follows: Paying tuition using aid resources does not significantly change student cumulative GPA (though it was seen previously to effect semestrial GPA). Overtime, nearly two thirds (64.5%) of individual student semestrial tuition was paid partially or in full using aid, and the common occurrence of students receiving some aid may explain this insignificance. The percentage of student tuition paid with aid has insignificant influence on cumulative GPA, as well as the cumulative number of semesters in which a student receives Pell or TAP. While Pell or TAP recipient status may indicate a higher level of external constraints, it was associated with a short term (semestrial) change and not a cumulative change to student's GPA. Paying tuition through self-resources has a positive association with cumulative GPA and was correlated with an expected increase of 0.049 GPA points per semester.

Overall, the correlation between the number of student peer-associations and cumulative GPA is limited to higher intensity peer-association (peers who met four or more times cumulatively). While peer association had an impact on the immediate semestrial GPA at all levels and associations, the cumulative influence observed is limited

Student Cumulative Credits Earned

One of the underlying hypothesis ($H2_d$) asserts that peer association is positively associated with student's cumulative credits earned. H2d questions if the correlation of student peer association is cumulative, and if it remains over time (even at times when the student does not have direct exposure to peers). What will be seen is that cumulative peer association is positively associated with cumulative credits earned. The correlation remains strong for all peer association intensity levels, and peer association remains stable in the presence of additional behavioral, academic, demographic and financial covariates.

The results of student peer-association on student cumulative credits earned are reported in Table 37.

Table 37

Students' Cumulative Credits Earned by Peer Association, Additional Behaviors, Academic, Demographic, and Financial Indicators (Models 1 through 5)

Y=Cumulative Credits Earned	Model 1	Model 2	Model 3	Model 4	Model 5
Behavioral Indicators					
Cumulative Peer Associations (Encountered 2 times)	0.395 *** (0.013)	0.275 *** (0.012)	-0.007 (0.007)	-0.008 (0.007)	-0.008 (0.007)
Cumulative Peer Associations (Encountered 3 times)	0.566 *** (0.023)	0.428 *** (0.022)	0.061 *** (0.012)	0.061 *** (0.012)	0.058 *** (0.012)
Cumulative Peer Associations (Encountered 4+ times)	0.451 *** (0.016)	0.334 *** (0.015)	0.027 ** (0.009)	0.026 ** (0.009)	0.022 * (0.009)
Cumulative Late Registration		-3.050 *** (0.214)	-0.957 *** (0.116)	-0.926 *** (0.117)	-0.871 *** (0.116)
Change Schedule		0.160 (0.100)	-1.717 *** (0.056)	-1.721 *** (0.056)	-1.683 *** (0.056)
Cumulative Return from Stopout		-13.403 *** (0.283)	-0.322 † (0.191)	-0.321 † (0.191)	-0.562 ** (0.196)
Academic Indicators					
Ever Need Writing (1st day of semester)			-0.225 † (0.133)	-0.220 (0.135)	-0.187 (0.135)
Ever Need Math (1st day of semester)			0.180 (0.132)	0.154 (0.136)	0.155 (0.136)
Total Credits to date			0.039 *** (0.002)	0.039 *** (0.002)	0.042 *** (0.002)
Cumulative Credits attempted			0.754 *** (0.009)	0.755 *** (0.009)	0.750 *** (0.009)
Cumulative Number of Classes			0.163 *** (0.027)	0.162 *** (0.027)	0.219 *** (0.028)
Cumulative Semesters Enrolled as part time			-0.232 *** (0.069)	-0.235 *** (0.069)	-0.102 (0.072)
Demographic Indicators					
Age				0.018 * (0.009)	0.016 † (0.009)
Male				-0.372 ** (0.127)	-0.375 ** (0.127)
Race/Ethnic Profile (Reference=White)					
Int. Student				0.140 (0.167)	0.091 (0.17)
Asian				-0.071 (0.235)	-0.024 (0.235)
Hispanic				0.103 (0.202)	0.046 (0.204)
African Americans				-0.129 (0.226)	-0.204 (0.228)
Other				-0.278 (0.215)	-0.311 (0.215)
Financial Indicators					
Cum. Semesters Paid with Aid (Any Portion of Tuition)					0.097 *** (0.029)
Cum. Semesters Paid with Self (Any Portion of Tuition)					-0.012 (0.063)
Cumulative Percent Tuition Paid using Aid					0.546 *** (0.15)
Cumulative Pell or Tap recipient					-0.791 *** (0.078)
Time (Semesters)	5.632 *** (0.084)	6.915 *** (0.089)	-0.164 † (0.088)	-0.175* (0.088)	0.017 (0.092)
Constant	-2.369 *** (0.147)	-2.911 *** (0.138)	-0.130 (0.104)	-0.355 * (0.29)	-0.718 * (0.302)

Note: Coefficients (and standard errors) reported, †p<0.1, *p≤0.05, **p≤0.01, ***p≤0.001

The relationship between student peer association and student's cumulative credits earned is consistent with those observed at the semestrial level. The influence of each peer association (encountered 2 times) is equivalent to an increase of 0.395 in a student's cumulative credits earned compared with a soloist student ($p \leq 0.001$). Similarly, the influence of student peer association at a higher exposure (encountered three times, and four or more times) is positive at 0.566 and 0.451 additional credits, respectively ($p \leq 0.001$). For each semester the student enrolls, the number of cumulative credits earned is expected to increase by 5.632 credits ($p \leq 0.001$).

In Model 2 the relative strength of peer association remains once coefficients for late registration, schedule changes and return from stopout are considered. Cumulative late registration behavior is associated with a decline of 3.050 in credits earned for each incident of late registration. A student returning from a stopout is expected have 13.40 fewer total credits earned as compared with a student that never stopped out. The influence of cumulative schedule changes on students' cumulative credits earned are insignificant (though schedule changes did influence student individual semester credits earned).

The correlation of student peer association changes in the presence of additional academic covariates (Model 3). Peer associations (encountered three, and four or more times) increases students' expected credits earned by 0.061 ($p \leq 0.001$) and 0.027 ($p = 0.002$) for each additional peer association, respectively. Having peer associations that were encountered twice had an insignificant influence on students' cumulative credits earned. The association of student late registration becomes weaker though it remains significant: Each late registration event is associated with a decline of 0.926 credits earned ($p \leq 0.001$). Student's schedule changes are now associated with earning 1.757 fewer semester credits for each semester in which the student changed their initial schedule ($p \leq 0.001$). Once academic covariates are accounted for, the influence of student's return from a stopout, and semesters enrolled become borderline significant and minimal: students returning from stopout are expected to earn 0.322 fewer cumulative credits compared with students who never stopped out ($p=0.091$), and it is expected that over the study period students would earn 0.164 fewer credits for each semester enrolled ($p=0.063$). Other academic indicators, which may in part capture peer association, have significant correlations with student's cumulative credits earned. Cumulative number of classes attempted is associated with an increase of 0.163 cumulative credits earned ($p\leq0.001$); Cumulative credits attempted is associated with an increase of 0.754 credits for each additional credit attempted ($p\leq0.001$). This association is expected as the number of peers associations and peer encounters opportunities is related to the number of classes a student attempts, and students' credits earned are directly associated with credits attempted. For each semester the student enrolls as a part time student, the expected cumulative credits earned declines by 0.232 ($p\leq0.001$). Students who needed Writing or Math remediation upon entry did not vary significantly in cumulative credits earned, though a semestrial need for remediation reduced the number of credits earned for a given semester[5].

Once student-level demographic indicators are included (Model 4), a similar correlation between peer association and cumulative credits earned remains. The correlation of student late registration, schedule changes and return from stopout remains similar to the one previously measured. In terms of student demographics, age and gender are associated with changes to the cumulative number of credits earned: age is associated with a small increase in credits earned (0.018 per year, $p=0.039$), and male students are expected to earn 0.372 fewer credits than

[5] See table 35 (model 3) for the between students' semestrial credits earned and selected behavioral and academic indicators.

female students (p=0.003). Differing from the relationship found in semestrial credits earned, there were no significant differences in cumulative credits earned attributed to race/ethnic profile. Any differences in cumulative credits earned by race/ethnicity[6] were accounted for using peer association, behavioral, and academic covariates. The use of financial aid to pay for a student's tuition was associated with increased cumulative credits earned, with an expected increase of 0.097 cumulative credits earned for each semester the student receives aid ($p \leq 0.001$). An additional portion of financial aid's influence is revealed using the covariate "percent tuition paid using aid". When tuition is completely paid for with financial aid the expected increase to a student's credits earned is 0.546 ($z=3.630$, $p \leq 0.001$). Repetitive Pell or TAP recipient status is associated with 0.791 fewer credits earned for every semester a student receives aid. Self-financed tuition without aid had an insignificant influence on cumulative credits earned.

To summarize, the association between peer-associations and cumulative credits earned takes place at higher intensity levels of peer-association (peers encountered three times, and four or more times). This correlation has a trend similar to the influence of peer association on students' cumulative GPA. The correlation of late registration, schedule changes and stopout to cumulative credits earned is negative, and carries a lasting influence on student credits earned.

Student Odds of Retention or Graduation.

One of the underlying hypothesis ($H2_e$) in this study questions the relationship of peer association to student's retention and/or graduation. It hypothesized that the correlation between student peer association and student retention and/or graduation occurs at the both the semestrial

[6] The average number of credits earned in the study period was 28.63. Hispanics, African Americans and other minorities earned on average 26.58, 26.33 and 27.60 credits, respectively. Asians, Whites and Int'l students earned 33.91, 32.45 and 33.00 credits, respectively.

and cumulative level, and that the correlation of peer associations remains over time. What will be seen is that students with peer associations is positively correlated with persistence into the following semester (or graduation) compared with soloists. The correlation remains positive (though changes in strength take place) once academic, demographic and financial indicators are taken into account.

The results of the initial model of student peer-association on student odds of retention are reported in Table 38.

Table 38

Student's Semestrial Odds of Retention or Graduation by Peer Association, Additional Behaviors, Academic, Demographic, and Financial Indicators (Models 1 through 5)

Y= Odds of Retention (or Graduation)	Model 1	Model 2	Model 3	Model 4	Model 5
Behavioral Indicators					
Peer Associations (Encountered Twice)	1.184 *** (0.012)	1.168 *** (0.011)	1.020 † (0.011)	1.020 † (0.011)	1.020 † (0.011)
Peer Associations (Encountered 3 times)	1.074 *** (0.018)	1.069 *** (0.017)	1.004 (0.017)	1.001 (0.016)	1.001 (0.016)
Peer Associations (Encountered 4+ times)	1.129 *** (0.016)	1.119 *** (0.015)	1.015 (0.015)	1.010 (0.015)	1.014 (0.015)
Late Registration		0.555 *** (0.044)	0.649 *** (0.053)	0.675 *** (0.055)	0.729 *** (0.06)
Change Schedule		0.787 *** (0.035)	0.692 *** (0.032)	0.711 *** (0.033)	0.707 *** (0.033)
Return from Stopout		0.568 *** (0.057)	0.718 *** (0.073)	0.716 *** (0.073)	0.717 *** (0.073)
Academic Indicators					
Need Writing (1st day of semester)			0.851 ** (0.052)	0.869 * (0.052)	0.878 * (0.053)
Need Math (1st day of semester)			0.686 *** (0.041)	0.719 *** (0.042)	0.732 *** (0.043)
Total Credits Earned (1st day of semester)			1.014 *** (0.001)	1.014 *** (0.001)	1.017 *** (0.002)
Credits attempted in the semester			0.984 * (0.008)	0.985 * (0.008)	0.977 ** (0.008)
Number of Classes in the semester			1.457 *** (0.043)	1.474 *** (0.044)	1.459 *** (0.043)
Part Time in the semester			0.893 (0.069)	0.916 (0.071)	0.828 * (0.065)
Demographic Indicators					
Age				1.007 * (0.003)	1.006 † (0.003)
Male				0.886 * (0.043)	0.887 * (0.042)
Race/Ethnic Profile (Reference=White)					
Int. Student				1.451 *** (0.122)	1.345 *** (0.12)
Asian				0.819 * (0.075)	0.812 * (0.074)
Hispanic				0.712 *** (0.056)	0.733 *** (0.058)
African Americans				0.762 ** (0.067)	0.792 ** (0.069)
Other				0.697 *** (0.058)	0.701 *** (0.058)
Financial Indicators					
Paid with Aid (Any Portion of Tuition)					2.141 *** (0.204)
Paid with Self (Any Portion of Tuition)					1.208 * (0.091)
Percent Tuition Paid using Aid					0.508 *** (0.066)
Pell or Tap recipient					0.925 (0.054)
Time (Semesters)	0.930 *** (0.009)	0.964 *** (0.011)	0.931 *** (0.013)	0.936 *** (0.013)	0.932 *** (0.013)
Constant	2.465 *** (0.116)	2.769 *** (0.139)	1.149 (0.135)	1.158 ** (0.19)	1.095 (0.202)

Note: Coefficients (and standard errors) reported, †p<0.1, *p≤0.05, **p≤0.01, ***p≤0.001

Students who established peer association were more likely to retain into the following semester compared with soloists. The correlation of peer associations and retention varied by intensity level: Peers who were encountered twice increased the likelihood of retention by 1.184 ($p \leq 0.001$); peers who were encountered three, or four or more times increased the likelihood of retention by 1.074 and 1.129, respectively. In the presence of additional behavioral covariates (model 2), both peer association and other behavioral indicators showed significant correlation to students' retention. Late registration for a given semester is associated with a reduction in persistence into the following semester (OR=0.555, $p \leq 0.001$). Student schedule changes also has an adverse effect on the likelihood of retention (OR=0.787, $p \leq 0.001$). A return from a stopout in a given semester is also associated with a decline in persistence into the following semester (OR=0.568, $p \leq 0.001$).

Similar to the analysis of students' earned GPA and credits earned, it is possible that the correlation between peer association and students' retention is partially explained using student-level academic covariates. Indeed, once student academic covariates are included (model 3), peer association remains positively associated with retention only for peers who were encountered twice. The association with peers encountered at higher intensity levels (more than twice) became insignificant. Part of the previously measured correlation between peer associations and students' retention is attributed to the number of classes a student attempts in a given semester, and to the total number of credits earned. Both variables had a positive influence on student retention (OR=1.457 and 1.014, respectively ($p \leq 0.001$)). In addition, this change may be partially attributed to having fewer students who maintain peer associations at high intensity levels in a given semester, though they might maintain such peer associations cumulatively[7].

[7] Table 39 revisits the model using cumulative indicators.

Students needing remedial Writing or remedial Math are less likely to be retained (OR=0.851 and OR=0.686, respectively) compared with those who are at college level. The influence of credits attempted per semester is small, though significant (OR=0.984, p=0.043).

The inclusion of demographic indicators (Model 4) does not change the correlation of peer association, behavioral or academic indicators previously measured. Peer association remains positively correlated (though borderline significant) only at the lowest intensity level. Late registration, schedule changes and stopout continues to penalize students' retention. In terms of student demographics, males and minority students are less likely to persist into the following semester compared with females and white (non-Hispanic students). Males have decreased odds (OR =0.886, p=0.012) of retention compared with females. Asian, African American, Hispanics and other minorities have 0.819, 0.762 and 0.712, and 0.697 times the odds of retaining into the following semester compared with whites ($p \leq 0.029$). International students are about nearly 1.5 times more likely to retain compared with white, non-Hispanic students (OR=1.451, $p \leq 0.001$). The likelihood of retention increases slightly with age (OR=1.007 per year, $p \leq 0.042$). This increase associated with age, though significant, does not carry a notable change to the likelihood of retention.

The addition of financial covariates in Model 5 does not change the association of previously measured behavioral covariates with student's retention. Using financial aid to pay for any part of the student's tuition more than doubles the likelihood of a student persisting to the following semester (OR=2.141, $p \leq 0.001$). However, because paying all tuition using financial aid reduces the likelihood of persistence on to the following semester (OR=0.508, $p \leq 0.001$), the influence of financial aid is also related to the proportion of tuition paid by aid. Pell or TAP recipient status (which was associated with lower student GPA and credits earned rate) does not influence student's retention, and students who self-fund any part of their tuition using non-aid resources are more likely to retain into the following semester (OR=1.208, p=0.013). Once financial indicators were included, the influence of part time status reduces the likelihood of student persisting on to the following semester (OR=0.789, p =0.002), and the number of credits attempted in the semester slightly reduces the likelihood of retention (OR=0.977, p =0.003). While the influence of credits attempted is significant, it is offset by the number of classes attempted in the semester (OR=1.459, $p \leq 0.001$). The influence of other demographic variables remains unchanged (though the influence of student age became borderline significant, p=0.088).

Overall, the models reveal that for a given semester, student behavioral indicators are associated with student retention into future semesters. However, peer associations' intensity level, which influenced GPA and credits earned, was not associated with the likelihood of semestrial retention. It is possible that the reason for the limitation on peer association intensity is attributed to the limited opportunity to establish high peer association intensity in such a short span of time as one semester. As students are allowed more time (semesters) to establish peer associations, the influence on retention and graduation may vary from the semestrial results.

Table 39 replicates the semestrial based peer associations analysis using cumulative covariates at each point in time, taking into account the student's cumulative peer associations, as well as the naturally occurring changes in each student's behavioral, academic, demographic and financial covariates at each point in time.

Table 39

Student's Cumulative Odds of Retention or Graduation by Peer Association, Additional Behaviors, Academic, Demographic, and Financial Indicators (Models 1 through 5)

Y= Odds of Retention (or Graduation)	Model 1	Model 2	Model 3	Model 4	Model 5
Behavioral Indicators					
Cumulative Peer Associations (Encountered 2 times)	1.048 ***	1.048 ***	1.013 *	1.014 *	1.013 *
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Cumulative Peer Associations (Encountered 3 times)	1.076 ***	1.075 ***	1.040 ***	1.036 ***	1.035 ***
	(0.008)	(0.008)	(0.009)	(0.008)	(0.009)
Cumulative Peer Associations (Encountered 4+ times)	1.053 ***	1.052 ***	1.022 ***	1.020 ***	1.020 ***
	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)
Cumulative Late Registration		0.633 ***	0.604 ***	0.645 ***	0.640 ***
		(0.036)	(0.038)	(0.039)	(0.04)
Cumulative Change Schedule		0.862 ***	0.744 ***	0.765 ***	0.753 ***
		(0.021)	(0.024)	(0.024)	(0.024)
Cumulative Return from Stopout		0.893	1.745 ***	1.563 **	1.912 ***
		(0.096)	(0.297)	(0.251)	(0.333)
Academic Indicators					
Ever Need Writing (1st day of semester)			1.061	1.079	1.080
			(0.068)	(0.066)	(0.068)
Ever Need Math (1st day of semester)			0.691 ***	0.738 ***	0.739 ***
			(0.043)	(0.045)	(0.047)
Total Credits to date			0.991 ***	0.990 ***	0.991 ***
			(0.002)	(0.002)	(0.002)
Cumulative Credits attempted			1.011 *	1.012 **	1.009 *
			(0.004)	(0.004)	(0.004)
Cumulative Number of Classes			1.077 ***	1.075 ***	1.095 ***
			(0.015)	(0.014)	(0.016)
Cumulative Semesters Enrolled as part time			0.909 *	0.930 *	0.850 ***
			(0.035)	(0.034)	(0.037)
Demographic Indicators					
Age				1.009 *	1.009 *
				(0.004)	(0.004)
Male				0.854 **	0.843 **
				(0.046)	(0.048)
Race/Ethnic Profile (Reference=White)					
Int. Student				1.400 ***	1.131
				(0.13)	(0.113)
Asian				0.909	0.918
				(0.096)	(0.100)
Hispanic				0.732 ***	0.767 **
				(0.066)	(0.072)
African Americans				0.798 *	0.839 †
				(0.079)	(0.087)
Other				0.783 **	0.799 *
				(0.075)	(0.079)
Financial Indicators					
Cum. Semesters Paid with Aid (Any Portion of Tuition)					1.050 *
					(0.024)
Cum. Semesters Paid with Self (Any Portion of Tuition)					1.099 **
					(0.035)
Cumulative Percent Tuition Paid using Aid					1.055
					(0.101)
Cumulative Pell or Tap recipient					0.870 ***
					(0.021)
Time (Semesters)	0.851 ***	0.909 ***	0.746 ***	0.755 ***	0.732 ***
	(0.01)	(0.02)	(0.034)	(0.032)	(0.034)
Constant	2.769 ***	2.965 ***	3.812 ***	3.529 ***	3.408 ***
	(0.128)	(0.168)	(0.298)	(0.524)	(0.547)

Note: Coefficients (and standard errors) reported, †p<0.1, *p≤0.05, **p≤0.01, ***p≤0.001

The correlation between students' cumulative peer association and student's likelihood of retention or graduation is positive and significant. Students' cumulative peer association changes the odds of student retention by 1.013 to 1.035 for each peer association established. Peer associations at higher intensity levels (encountered three times, and encountered four or more times) have greater influence on students retention compared with cumulative peer associations that were encountered only twice: 1.305 and 1.020 for each peer encountered three and four or more times, respectively. The influence of late registration and schedule changes on cumulative retention remains negative, and it is associated with changes of 0.640 and 0.753 in the cumulative odds of retention, respectively ($p \leq 0.001$). The association between a return from stopout and students' retention over time was associated with an increase in the odds of retention ($OR=1.912$, $p \leq 0.001$). The direction of the association is opposite to the one measured at the semestrial level. It is possible that return from stopout signals higher commitment to the academic path in the presence of external obstacles (e.g. work or family commitment), or that it indicates that the student addressed previous constraints to their schedule. Over the study period this outcome effected over one fifth (22.2%) of students, who experienced at least one stopout.

The influence of student cumulative academic indicators slightly differed from those measured at the semestrial level. Requiring Writing remediation is now insignificant, and students who entered college needing Math remediation are less likely to retain to a future semester ($OR=0.739$, $p \leq 0.001$), though the influence is weaker than the one measured at the semestrial level (reported between $OR=0.686$ and $OR=0.732$ in models 3, 4 and 5). This variation is expected, as the cumulative variable reflects on students' historical preparation which will change if a student exited remedial mathematics, while the semestrial indicator marked the status of the student at the beginning of a given semester. While significant, the influence of credits attempted to date on students' retention is minor (OR=0.991, $p \leq 0.001$) and is offset by the number of classes taken to date (OR=1.095, $p \leq 0.001$). For each semester the student enrolls as part time, the likelihood of retention declines by 0.850 ($p \leq 0.001$).

Males and minority students are less likely to retain to the following semester when compared with females and white (non-Hispanic students). Males have 0.843 times the odds of retaining compared with females ($p = 0.003$). African Americans, Hispanics and other minorities have 0.767 and 0.839, and 0.799 times, respectively, the odds of retaining into the following semester compared with whites. The changes in the odds for African Americans is borderline significant ($p = 0.091$). There are no significant differences in the odds of cumulative retention for international students compared with whites, this change from the semestrial models may be attributed to changes in visa status over time. During the study period 22% of international students have changed the type of visa they possessed.

In terms of financial aid, each semester in which a student used aid-funding to pay any part of their tuition had increased by 1.05 their odds of retention. The number of semesters for which a student received Pell or TAP aid is associated with lower student's retention (OR=0.870, $p \leq 0.001$). The cumulative influence of semesters enrolled is stronger (OR=0.732, $p \leq 0.001$) than that measured in a semester-based model. This is to be expected, as the number of semesters enrolled increases, the likelihood of an external constraint to influence a student's ability to enroll or complete their studies increases.

Overall, the cumulative models reveal that cumulative behavioral indictors, such as peer association, late registration, schedule changes and return from stopout, may be associated with student retention and graduation rates. The models reveal that for a given semester, student cumulative behavioral indicators are associated with student retention into future semesters. Peer associations' intensity level, which influenced GPA and credits earned, is associated with the likelihood of overall (long-term) retention (but it does not have a significant influence on students' semestrial retention).

CHAPTER V

In this chapter, the study's hypotheses are reviewed, and conclusions are drawn from the analysis of the study's findings. Then the common themes between this study and other studies pertaining to peer association and behavioral indicators are discussed, as well as the contribution of this study to the literature on community colleges. It further discusses the implications of the findings, recommendations related to informing policy makers, and the benefits of using student-level transactional data techniques to extract knowledge from administrative student records. The chapter concludes by considering the limitations of this study and proposing other possibilities for continued research on peer association and also other uses of student-level transactional data at community colleges.

Discussion of the Findings

This study examined various aspects of student behavior and also demonstrated that records management systems can be mined to extract longitudinal student-level transaction data. Extracting data in this way provided the study with a proxy for behavioral indicators such as: peer association, late registration, and schedule change. These behaviors were correlated with various longitudinal academic outcomes, such as: student GPA, credits earned, and retention (graduation). One of the main findings from the results of this study was to demonstrate the ability to identify and measure peer association directly from transactional files. For example, this technique would certainly enhance the findings of previous studies which were limited to using only short-term or single point in time data, such as cohort membership, dormitory roommate, or squad assignment as evidence of peer association. Whereas other studies assumed that such relationships would endure over time, the structure of this study allowed for the accumulation of a wide range of long-term fact-based data elements which were verified during the iteration of each repeated measurement. The benefit of such a large and complete data set established concrete evidence enabling rigorous analysis which yielded correlations between student behavior and student academic outcomes, both at the semestrial level, and cumulatively.

This study uses student-level transactional data (live and historical data) stored in information systems supporting student learning or administrative records to focus on measures of student peer association in the natural setting of a large open admission, urban, commuter community college. Those types of records can be used to generate a time-dependent and accurate profile of the student behaviors (Baepler & Murdoch, 2010; Ferguson, 2012; Cristobal Romero & Ventura, 2013). Until now, the majority of literature on peer effects in postsecondary education has focused on traditional age students attending four-year residential colleges. The studies at community colleges related to peer effects concentrated on the relation of outcomes associated with learning communities (Richburg-Hayes, Visher, & Bloom, 2008; Rocconi, 2011; Weiss et al., 2014). Learning community studies which used randomized trials (Richburg-Hayes et al., 2008; Visher et al., 2010; Weiss et al., 2014) assumed that the peer structure inherent to learning communities was tied directly to peer association, however the existence of this association was not substantiated by longitudinal student registration records. In other words, previous studies did not examine whether the learning communities settings indeed created longstanding peer associations that were maintained across other curricular activities as well as over time.

The ability to verify the accuracy of student data, and the point in time at which that data is taken, is critical. This study found that nearly three quarters (73.2%) of students made modifications to their initial course registrations at least once during the study period, while an average of two fifths (41.2%) of the population change their initial schedule each semester.

These changes, unrecorded in other studies For that reason, constructs such as learning communities, which are intended to encourage peer associations on a large scale may not actually be as effective as reported (J. Smith & Stange, 2015). Other studies exploring peer association at commuter institutions relied on indirect measures of both association and peer quality, such as student high school attendance, residence or PSAT scores. For students attending community colleges, those measures are likely to be confounded for several reasons, such as: being of non-traditional age, delay to postsecondary enrollment after high school, loss of association with their high school fellows, potential residential changes since high school and during college attendance, lack of PSAT or other standardized tests scores, and didn't graduate high school but earned a GED.

The Influence of Peer Association on Student Outcomes

This study found that student peer association can be documented and monitored in the natural setting of a commuter community college. Analysis revealed that peer association levels were correlated with a differentiation in students' academic outcomes at a single point in time (semestrial level) and cumulatively. Table 40 summarizes the main contributions student-level transactional data analysis provided to the measurement of curricular student peer association. Also included are key findings pertaining to the correlation between student peer association and academic outcomes, followed by suggestions for establishing new practices or policy in response to the new findings.

Contributions of transactional student-level data analysis to this study	(a) Sampling all students eliminates the need for randomized trials to identify study subjects and captures the dynamics of all curricular related exposure to peer structure (b) Measurement of curricular peer association can be taken at a point in time, or cumulatively (c) Using transactional student-level records provides higher data accuracy and fidelity, allowing identification of actual proximity to curricular peers that cannot be achieved through traditional methods of data collection
Findings that are substantially new (different)	(I) Despite variation in peer intensity, and number of semesters studying without peers, soloists differ (typically underperform academically) compared to students with peers. (a) Who are the soloists? Soloists are more likely among : older students, Intl' students African Americans, college transfers, , and students who enter fully prepared for college level coursework (no remedial coursework) (b) What proportion of the student body are soloists? Roughly two thirds (64.2%) of the students were soloists at some point in time. Long term 27% of the students were soloists for at least half of their enrollment period, and 10.1% of the students were always soloists. Previous studies did not document peer association across multiple semesters (c) Only 30%, or less, of single course takers carry peer associations from previous semesters Single course takers are more likely to be soloists at a given point in time, and over time, compared with multiple course takers. (II) The relation to other student-level behaviors (a) Students with the fewest peers (and soloists) are more likely to practice late registration, change their schedule once classes begin and skip new student orientation. Low levels of student peer association is associated with early departure from college. (b) Soloists are more likely to take lower course load levels in a given semester, and over time. (III) The association to academic outcomes (a) The longer a student is a soloist, the fewer credits they attempt, and earn; their GPA is lower in the semester, and also over time.
How should colleges react based on the new findings?	**(I) Planning:** (a) In addition to learning communities, colleges should consider additional ways to encourage student peer associations. Even small student networks contribute to improved student outcomes, and may be simpler to implement than large scale learning communities. (b) Although no standard exists between colleges for comparing transactional data nationally individual colleges, and some college systems, can benefit realty from transactional data to evaluate students' behavior on a real time basis **(II) Outreach and advisement:** (a) Colleges should utilize the information available from transactional records to recognize soloists (and students with small/'decreasing' networks) and target an outreach/advisement to those students to aid expand their peer networks (b) Colleges should reexamine students' at-risk classification patterns not only on the basis of their pre-college or first semester data, but also across time. Using peer association data as it evolves over time will allow colleges to better align the available (limited) resources for student support with students in need. (c) College can utilize 'soloist' status as a risk indictor for early college departure. **(III) Registration Processes and Monitoring of Student Academic Progress:** (a) Colleges may utilize peer association information to 'populate' the student suggested schedules with classes that have potential curricular peers. (b) Once identified, soloist students may be monitored for other negative academic markers, in particular: late registration, schedule changes, and reduction in course load. Soloist students may be further monitored and outreached for academic performance and support

Table 40. *Summary of the Findings for Student Peer Association*

The current study's ability to measure peer association directly builds on the work of previous studies (e.g. Lyle (2007); Winston and Zimmerman (2004)), where peer association was based on the random assignment of students as dorm roommates or squad members. The authors of those studies were able to measure the effect of roommates, but at the same time identified limitations of the random assignment methodology. Superficially, the assumption of student peer association through assignment (without verifying its existence) might result in associations being estimated when they did not actually exist. Similarly, Stinebrickner and Stinebrickner (2006) acknowledged that previous peer studies focused on outcomes that "take place relatively quickly after a student receives the "treatment" of being assigned a roommate" (pp.1437), and for that reason may have a bias in the estimation of peer association (either over or under estimate the peer's association influence). Furthermore, Stinebrickner notes, "given the difficulty of credibly identifying peer effects in education, much of the literature on peer effects in education has had the goal of trying to find compelling evidence that peer effects can matter" (pp.1452). The primary limitation in all that research was an inability to identify curricular associations attributed to the study's random assignment or peers.

In this study, curricular peer associations were measured directly from the student record, and assessed both at the semestrial level and cumulatively. The student-level transactional data enabled the examination of each student's record to generate a virtually complete mapping of curricular peer association, at an accuracy level which was not feasible in previous studies utilizing survey instruments or group assignment dependent). Furthermore, transactional data allowed to minimize the influence of unmeasured covariates on students' outcomes that may be reflected through peer association (but peer association was not determined through random assignment)

As previous peer association studies have noted, there is a challenge identifying causal relationships in peer-effect studies with regard to pre-determined characteristics (unobserved variables), in addition to separating peer-effects from other confounding covariates (e.g. Manski (1993), Sacerdote (2001, 2011), and Hanushek et al. (2003)). The challenge other studies shared in common had to do with the limited set of covariates available, and consequently, the ability to control for unobserved variables was limited.

Through the use of administrative records and student-level transactional data, the ex post facto structure of this study did not rely on randomization to construct student peer groups or the exposure to peers, but instead identified naturally occurring peer relationships directly through observation of an entire student population. The measurement of student peer groups was conducted for a complete, longitudinal and larger sample compared with other studies that evaluated naturally occurring peer associations (e.g. Winston and Zimmerman (2004)). The inherent accuracy this method provides in terms of identifying the actual proximity among specific peers cannot be achieved through other traditional methods of data collection, and certainly not practical at the scale achieved in this study. When compared to previous studies, this study accounted for academic, demographic and financial variables at baseline, and went on to perform updates at key points in time each semester to reflect changes over time. Those measurements accounted for common unobserved variables inherent to academic, demographic and financial attributes that might influence student outcomes correlated with curricular peer association. Furthermore, through an exhaustive revisit of the covariates each semester, this study provides enhanced control to the ‘order of time’ influence. Simply put, all covariates were re-measured at the beginning of each semester allowing to better evaluate the influence of student peer association for a given semester, and cumulatively, while controlling for other known covariates.

In this study student peer-association was found to be correlated with academic outcomes by measuring dependent covariates, ex post facto, directly from the transactional records and student academic transcript records. These correlations were associated with both the number of peer associations a student established, and with the intensity (repetitiveness) of peer encounters. The measurements of the number of peers and their intensity is novel to studies of peer association in community colleges, and may be relevant for open admission commuter institutions in general. While previous post-secondary studies examined residential, highly selective four-year institutional data (Foster, 2006; Sacerdote, 2001, 2011; Zimmerman, 2003), this study evaluates student peer association in a commuter setting of an open admission community college. This distinction is crucial as it allows similar institutions to evaluate student peer-association in the natural settings of the college, and measures students' peer-association as they develop (within each semester and over time)

At the semestrial level, the number and intensity of student peer associations in this study was found to be correlated with student semester GPA and semester credits earned. This correlation showed the most pronounced difference among soloist students, who earned lower GPA and lower credits in a given semester when compared with students who established many, peer associations. Interestingly, this difference remained in place, although to a lesser degree, between soloists and students who established only a single peer association.

Starting with the first semester, and throughout study period (13 consecutive semesters), the influence of peer association was statistically significant and positive at all levels of peer association intensity. Surprisingly, the influence of peer associations intensity levels on semester credits earned declined with increased intensity[8]. It is possible that part of the peer-associations intensity level was captured within the number of courses and credits attempted. These variables reflect on the student's course load and, in turn, will also influence the opportunity for a student to establish semestrial peer associations. Semestrial peers increased the likelihood of student's persistence (or graduation) into the subsequent semester, but the influence in a given semester was significant only for the lowest level of peer association intensity (encountered peer only twice). It is likely from the findings that peer association intensity at the semestrial level was associated with, and limited by, student course load. Perhaps the most important finding was that the cumulative effect of peer associations related to retention (graduation) over the course of a student's enrollment was positive and significant at all intensity levels.

Similarly, this study shows that soloist students had lower cumulative GPA and earned fewer cumulative credits compared with students who established some peer associations. This study shows that student cumulative GPA, credits earned and retention rates are correlated with the number of student peer associations, and by the level of peer association intensity. The influence on cumulative GPA is limited to the highest level of peer association intensity (four or more repeated peer encounters). Similarly, for all but the lowest level of peer association intensity (encountered twice), the influence of peer association on cumulative credits earned was significant. The influence of peer association on student's persistence (cumulative retention) or graduation was positive and significant. As the number of peers and the intensity of peer association increased, the odds of a student persisting increased as well. Despite variance in peer

---

[8] The standardized scores for semester credits earned started at 6.77 for single peer associations (peer encountered only two times) and declined to 2.97 for peer associations encountered four or more times), though the influence of the number of peer associations remained positive and significant (see also Table 35).

association levels (number and intensity) attributed to students' academic and demographic covariates, the influence of peer association was consistent when all other factors were held constant.

The influence of student peer association measured in the study supports the findings of Stinebrickner and Stinebrickner (2006) - when evaluating the role played by student's curricular interaction they reported that curricular interaction contributed more to peer effect than that of either roommate or other dormitory interactions. The findings in this study demonstrated that the influence of peer association is significant and that the correlation remains over time when academic, demographic and financial covariates are taken into account. The findings further support the conclusions of previous studies that reported positive peer effects in learning communities on students at non-residential institutions (Chesebro et al., 1999, June; Engstrom & Tinto, 2008; Shapiro & Levine, 1999; Stassen, 2003; Vincent Tinto, 2003; Zhao & Kuh, 2004). Unlike other studies that assumed peer-association influenced student outcomes due to the structure of learning communities, this study measured peer association directly and helps strengthen the argument for the influence of peer association. This study controls for multiple covariates over the study period that may reflect on student peer associations, and on the measured outcomes. The measurement of academic, demographic and financial indicators from the administrative records allowed for rigorous evaluation of peer associations influence. Furthermore, the study's findings are aligned with those in the previous literature, suggesting that peer association is in fact associated with small yet positive effects on academic outcomes such as GPA (Foster, 2006; Sacerdote, 2001, 2011; Zimmerman, 2003). The implication is that peer influence, previously observed at four year residential institutions may also apply to open admissions, commuter community colleges.

The Correlation between Late Registration and Student Outcomes

This study found multiple student behavioral indicators which were associated with a differentiation in students' academic outcomes. Table 41 summarizes the main findings for the correlation between student late registration patterns and academic outcomes:

Capabilities of Transactional student-level Level Data	(I) Detailed capture of all transactions during a full registration cycle: (a) Track cohort and individual student late registration activity directly from administrative records (b) Sort late registration activity (i.e. course registrations) by precisely defined events common to all colleges (e.g. first day of registration, first day of classes, and last day of a semester) (c) Transactional data promotes a standard definition of late registration, easily verifiable and replicated across institutions. (II) Completeness of the records: (a) Transaction data analysis is a non-intrusive method for determining the individual late registration behavior of every student, either while it occurs or ex post facto. (b) Freed from the administrative burdens and low response rates of surveys and questionnaires, the researcher can expect data points on virtually all students by using the student transactional data, not just a subset of the students. (c) The fidelity of data elements is not vulnerable to misinterpretation by survey respondents. (III) Semestrial and longitudinal evaluation: (a) Transactional student-level data permits evaluation of late registration indicators over time, regardless of whether or not the student was enrolled continuously, or even if no longer enrolled. (b) The transactional data allows the researcher to establish different weights to recent behavioral indicators compared with historical behaviors of the subject (e.g. lateness in a recent semester may provide a better indication than lateness that took place in a previous year).
Findings that are substantially new (different)	(I) Standard Definition of Late Registration: This study developed a portable definition of late registration usable at any college. (II) Thorough Documentation // Magnitude Estimation: (a) This study documented the magnitude of late registration behavior, not only at college entry, but also across time (b) Nearly a 25% of students (718 out of 3,128) registered late at least once, and more than 10% have registered late multiple times. Previous studies had limited data and could not accurately estimate the longitudinal extent of late registration behavior. (II) Correlation to academic student outcomes (a) Late registration was correlated with immediate as well as long-term negative consequences to student academic outcomes. (b) Late registration was associated with an increase likelihood of a student being either a soloist, and/or having few peer associations compared to the general population of the study cohort (c) Late registrants attempted and earned fewer credits; had lower GPA in the semester they registered late, and also over time. (d) Late registrants had lower retention and graduation rates in the semester they registered late, and cumulatively (even in a later semesters where they registered on time). Students were 27% less likely retain or graduate in the semester they registered late, and 36% less likely to retain or graduate over time compared with on time registrants (IV) Capabilities: (a) Community Colleges can identify and measure registration patterns directly from the transactional file, in lieu of surveys (b) The measurement of student late registration can take place within a semester, and over time (c) Identify late registration patterns in individual students.

	(d) Correlation of negative student outcomes and late registration remained in place after accounting for academic, demographic and financial covariates. Those variables were measured directly from the administrative records and constantly updated to account for changes in the student profile over time.
How should colleges react based on the new findings?	**(I) Planning:** (a) Colleges have more evidence to support eliminating late registration. . (b) Community colleges can should consider collecting data from their transaction recording to better understand students' behavior on a real time basis, specifically regarding late registration **(II) Student Intervention** (a) Late registration behavior, in combination with other at-risk factors, are useful to signal that some sort of student intervention may be warranted. (b) Late Registration behavior, and other factors now available post-admission using the student-level transactional data, should be added when identifying at-risk students, targeting interventions, and developing communication strategies. **(III)Activities, Registration and Monitoring:** (a) Because late registration is associated with other negative behaviors (e.g. lower levels of peer association), colleges consider developing types of wrap-around services to address several issues at once. (b) Colleges may consider policies that abolish or discourage late registration

Table 41. *The Correlation between Late Registration Patterns and Student Outcomes: Capabilities of Student-Level Transactional Data Analysis, Central Findings and Potential Reactions for Community Colleges.*

The advent of transactional student-level data analyses used in this study produced novel findings by capturing all student-level transactions, over the study period (13 consecutive semesters) to reveal longitudinal evidence never before identified. Students demonstrating late registration behavior overwhelmingly had lower academic performance when compared to on time registrants. The analysis revealed negative consequences related to student academic performances took place at both the semestrial level and cumulatively. By comparison, previous studies sampled students for only one semester at a single institution by using a randomly selected samples (Angelo, 1990; Chilton, 1964; Diekhoff, 1992; Keck, 2007; Mannan & Preusz, 1976; Neighbors, 1996; Parks, 1974; Peterson, 1986; Street, 2000), and of these (Belcher & Patterson, 1990; Goodman, 2010; Sinclair Community College, 2005; Stein, 1984) went so far as to examine late registration using the entire student body. Only one study (Hale, 2011) evaluated the influence of late registration indicator on a students' cohort for a period that exceeded two years.

In contrast, this study measures students' late registration using existing, verifiable transaction level data elements which represent each day's registration activity. Unlike the information contained in student transcripts, the use of transaction data by this study avoided inaccuracies present in post-processed or 'backdated' information – the result of the common practice of clerical revisions. Transactional data provides a very accurate picture of student behavior, which provides time and date stamp to the student actions and allows for a precise measurement of late registration. Previous studies suffered a risk of underestimation of a student late-registration pattern, as the studies employed a 'frozen' dataset, which did not account for backdated information. Registration activities that took place on or after the first date of classes may have reflected dates prior to the semester. For similar reasons the accuracy of the measurement in this study is superior to studies that have used survey information to evaluate late registration behaviors, as students and administrators may report late registration differently (due to misinterpretation of the registration times, not remembering the registration activities taken, backdated registration activities, etc.)

The current study is the first to provide thorough documentation of late registration occurrence for an *entire student* cohort. Over the 13 consecutive semesters of this study nearly one in four students (718 out of 3,128 students) had practiced late registration at least once. The proportion of the student body that practices late registration throughout their enrollment period requires attention from community colleges administrators and faculty, especially in relation to its association with longitudinal negative academic outcomes. Specifically, late registration was associated with a decline in student's semester GPA and semester credits earned. The first instance of late registration was correlated with a lasting influence on student GPA and credits earned, and remained statistically significant overtime. While it was not entirely possible to rule out preexisting student characteristics at any given time, the study's process of re-evaluating every student profile each semester had the effect of differentiating innate versus unintentional behavior. In terms of student retention, late registration was found to be associated with a reduction in the odds of student persistence (or graduation) into the subsequent semester, and to an even greater extent so too was the effect of repetitive late registration behavior.

The findings of this study align with previous studies that showed late registration is predictive of earning lower GPA (Freer-Weiss, 2004; Hiller, 2005; Moore et al., 2007; Roueche & Roueche, 1993; Smith et al., 2002; Summers, 2003); and that students were less likely to persist in college in subsequent semesters (Freer-Weiss, 2004; Johnston, 2006; Smith et al., 2002; Summers, 2000). While other studies varied in their definition of late registrations, this study used a strict definition for late registration that allows replication across institutions[9]. In addition, this study measured late registration not only at entry, but also across time. It evaluated the influence of both semestrial and cumulative late registration indicators, at any point in time during the student's enrollment.

The current study contributes to the literature by demonstrating that late registration can be measured directly from the transactional record without conducting a collection via survey that is subject both to survey fatigue and to concerns with data completion and data accuracy. The transactional data ensures virtually complete data for all students in the cohort, while adding no load to an existing data collection mechanisms that use direct surveying. The current study further addressed a concern regarding the lack of a standardized definition of late registration, and employed a late registration definition which is replicable across all institutions. Second, this study builds on previous studies, adding to the literature the ability to measure the longitudinal influence of late registration beyond the semester or point-in-time in which it took place. Previous studies either made no attempt, or were limited in the evaluation of the longitudinal influence of late registration - the current study addresses this limitation through an evaluation of both point in time and cumulative (repetitiveness) late registration. Controlling for additional academic, demographic and financial covariates this study enables to further 'control for the influence of late registration as in indicator for students' underlying characteristics, or as an indication for external constraints that may influence student outcomes.

The Correlation between Student-Initiated Schedule Changes and Student Outcomes

Table 42 summarizes the main findings for the correlation between student schedule changes patterns and academic outcomes.

---

[9] Late registration is defined as having the first registration activity for the semester taking place on or after the first day of classes. See pp. 12-13 for a discussion of the multiple definitions of student late registration.

Capabilities of Student-Level Transactional Level Data Analysis	(I) Detailed capture of all transactions during a full registration cycle: (a) Track cohort and individual student registration activity directly from administrative records (b) Sort schedule change activity by precisely defined events common to all colleges (e.g. first day of registration, first day of classes, and last day of a semester) (c) Transactional data analysis is able to gather a complete record of schedule change activity. Previous studies, even those that used administrative records, only considered the final outcome based on transcript data and missed the many transactions that might have occurred. (d) Transactional Data can recognize patterns of course schedule changes, such as ‘course shopping’ or ‘bulk shopping. (e) Using transactional data, this study was able to differentiate student initiated schedule changes from administrative course schedule modifications. (II) Completeness of the records: (a) Transactional Data is a non-intrusive method for determining the individual late registration behavior of every student, either while it occurs or ex post facto. (b) Freed from the administrative burdens and low response rates of surveys and questionnaires, the researcher can expect data points on virtually all students by using transactional data, not just a subset of a student population. (c) The fidelity of data elements is not vulnerable to misinterpretation by survey respondents (III) Semestrial and longitudinal evaluation: (a) Transactional Data permits evaluation of schedule change behavior over time, regardless of whether or not the student was enrolled continuously, or even if no longer enrolled. (b) Transactional Data allows the researcher to establish different weights to recent behavioral indicators compared with historical behaviors of the subject (e.g. schedule changes in a recent semester may provide a better indication than schedule changes that took place in a previous year).
Findings that are substantially new (different)	(I)Detailed mapping of student initiated schedule change: Previous studies looked at schedule changes as one of several variables of interest, and lacked a complete evaluation of its occurrence and 'source'. (II) This study developed what might be a more useful and perhaps universally comparable definition of schedule changes: student imitated change, taking place on or after the first day of classes. (III) Thorough Documentation // Magnitude Estimation: (a) This study documented the magnitude of its occurrence not only at college entry, but also across time (b) Nearly three quarters (73% of the students) changed their class schedule at least once during their studies. It differs from previous studies that surveyed student schedule changes for a single semester, and found the behavior to occur with only 15%-25% of students. (c) Schedule changes were correlated with immediate as well as long-term negative consequences to student academic outcomes. Previous studies were limited in the evaluation of longitudinal outcomes. (IV) The relation to other negative behaviors (a) this study found a correlation between schedule changes and the likelihood of a student being either a soloist, and/or having fewer peer associations when compared to the general population of the study cohort (V) The association to academic outcomes (a) Schedule changers attempted and earned fewer credits; had lower semester and cumulative GPA This study is the first to produce a longitudinal analysis of the negative influence associated with schedule change behavior. (b) Schedule changers had lower retention and graduation rates into the semester following schedule change, and cumulatively (even in a later semester where they registered on time). To date, this is the first study documenting a cumulative negative influence which extends beyond the semester/academic year level.

	(VI) Capabilities: (a) Community Colleges can identify and measure registration patterns directly from the transactional file, in lieu of surveys; completes a missing layer of information that is not available via student transcript or traditional administrative records (b) Measure schedule change behavior, such as course shopping and bulk shopping, each semester, and over time (c) Record schedule change transactions, as they occurred, for each individual student. This ability overcomes the limitation encountered when records are overwritten or modified by administrative processes. (d) The correlation of schedule changes to negative student outcomes remains in place once academic, demographic and financial covariates are accounted for. Those variables were measured directly from the administrative records and constantly updated to account for changes in the student profile over time.
How should colleges react based on the new findings?	**(I) Planning:** (a) Colleges should facilitate student course selection prior to the first day of classes in an effort to reduce the rate of schedule changes. (b) Schedule change data can be used to inform the course/section provisioning process. (b) Using transactional data gathered on a real time basis, students who excessively change schedules can be identified early and interventions can be provided. **(II) Outreach and advisement:** (a) Colleges should utilize the information available from transactional records to recognize those students with chronic schedule change behavior, and target outreach/advisement to them. Such outreach would increase course availability to all students, as schedule changers occupy space in a class they eventually will not take. (b) Colleges that reexamine students' at-risk classifications should consider adding schedule change as a risk indicator. (c) College can revise communication and services strategy to address constraints that lead to chronic changes (d) Colleges may consider means to improve student course selection, e.g. introductory class for each course prior to the semester, which would allow students to self-assess their fitness to the curriculum. **(III)Activities, Registration and Monitoring:** (a) Schedule Change is associated with other negative behaviors (e.g. lack of peer associations), therefore colleges may utilize this information to link students with other potential peers during the registration process (e.g. 'populate' the student suggested schedules with classes that have potential curricular peers.) (b) Additional advisement, and regular monitoring may be added to students who practice schedule changes regularly to ensure they remain on track towards a degree.

Table 42. *The Correlation between Schedule Changes and Student Outcomes: Capabilities of Transactional Student-Level Data Analysis, Central Findings and Potential Reactions for Community Colleges.*

This study contributes to the literature by demonstrating that student initiated schedule changes can be measured directly from the transactional record without conducting a collection via survey that is subject both to survey fatigue and to concerns with data completion and data accuracy. Using transactional data ensures a virtually completed data collection for all students in the cohort, while adding no load to an existing data collection mechanisms that use direct surveying. The current study further addressed a concern regarding the lack of a standardized definition of schedule changes, and employed a definition which is replicable across all institutions. Second, this study builds on the findings of the few previous studies which explored schedule changes as a factor in student academic progress, by adding to the literature the ability to measure the longitudinal influence of schedule changes beyond the semester or point-in-time in which it took place. Student-level transactional data allows mapping patterns of schedule changes by compiling a complete record of the changes capable of differentiating a student-initiated versus campus, academic department, or other administration initiated changes (e.g. class cancellation or campus schedule changes). Previous studies were limited in the documentation of schedule changes, and in the evaluation of longitudinal influence of schedule changes. The current study addresses these limitation through an evaluation of both point in time and cumulative (repetitiveness) schedule changes. Controlling for additional academic, demographic and financial covariates this study enables to further 'control for the influence of student initiated changes as in indicator for students' underlying characteristics, or, as an indication for external constraints that may influence student outcomes.

Similar to late registration behavior, student initiated schedule changes had a consistent, negative association with student academic performance at the semestrial level and cumulatively. Schedule change was correlated with lower levels in student's semester GPA and semester credits earned. The cumulative influence of schedule changes on student GPA and credits earned remained significant overtime, though the influence of schedule changes on student GPA was statistically more significant at the semestrial level compared with the cumulative influence. In terms of student retention, schedule changes were correlated with a reduction in the odds of student retention in to the subsequent semester, as well as with a reduction in the cumulative odds of student retention (or graduation).

This study is unique as it adds to the very few studies (e.g. Broadbent (1975) and Hagedorn et al. (2007)) that specifically targeted student schedule changes within student behavioral studies. Unlike those previous studies, which used student surveys to evaluate course changing behaviors, this study evaluated schedule changes directly from the transactional data. The transactional data employed in this study allowed for a virtually complete and accurate evaluation of student initiated schedule changes, which cannot be captured in a student survey or through an analysis of transcript data alone. For example, in BPS 2004/09 students' transcripts were examined to determine if a student changed a course following an institutional add/drop period. However, as the description of the BPS2004/09 report methodology reported, they could only identify official withdrawals - students change courses for many types of reasons that are not recorded on the student transcript. Among public, community college students who had no course changes reported on the transcript, 18% reported that they had changed a course schedule at least once during their studies (NCES, 2015). Transactional level data used in this study went further by excluding schedule changes that were not student initiated (i.e. college-level or administrative reasons such as class cancellation or time changes). This study also limits the evaluation of schedule changes to ones that take place in or after the semester starts. The findings further documented that student initiated schedule changes may be a negatively influencing indicator on student outcomes. Previous studies were not able to document the influence of schedule changes (or determined that its influence is neutral) because schedule changes were underreported, and subject to different interpretations of its definition. Students and administrators did not necessary report all (nor accurately) schedule changes due to incomplete surveying, and due to a potential misinterpretation of questions pertaining to schedule changes. This study measured the changes directly, and was able to document the influence of schedule changes on semestrial and longitudinal student outcomes. The negative influence of schedule changes is supported by Summers (2000) who documented a decline in student GPA for students practicing schedule changes, as well as a decline in student's persistence rate. This study's findings demonstrated that the practice of schedule changes occurs at a high rate throughout a student's enrollment. Over the study period, 41.2% of students had a schedule change in their first semester and 73.5% of all students in the study initiated a schedule change at least once throughout the study period. This pattern was not revealed in previous schedule change studies which focused on a single point in time and did not evaluate the students' practices over time. Neither was a longitudinal evaluation of schedule changes performed in previous studies evaluating schedule changes. Furthermore, previous studies relied heavily on student surveys or limited transcript data that rarely documents all schedule changes. The transactional data provides new insights with regard to the registration cycle, as it accounts for all changes (and their source) at the time of event occurrence. The high rate of schedule changes further illustrates the reasons for which course pairing or learning communities' settings may still face a challenge steering students to develop peer associations: while the initial schedule may ensure a student peer association, the common practice of schedule changes may reduce student peer association rates.

The influence of student stopout behavior on academic outcomes was mixed. Students returning from stopout earned higher semestrial GPA in the semester in which they returned, while the influence on number of credits earned was borderline significant, yet still positive. Interestingly, returning from stopout was associated with lower retention or graduation in the semesters following their initial return. Cumulatively, return from stopout was associated with higher cumulative GPA, and an increased likelihood of retention (or graduation) over time. This is not surprising, as students who return from stopout may have overcome some external obstacle(s) to their studies, as well developing a greater commitment to their degree path. Naturally, stopout behavior was associated with a reduction in the number of credits earned as the number of semesters a student enrolled declined. Previous studies were limited in their ability to study stopouts due to the shortness of the study, or were unable to match the student's departure and return record at the institution. Over the course of this study nearly one in five students who stopped out returned, and provides an approximation for a stopout rate that was not available in previous longitudinal studies of urban, public community colleges. While the majority of students remained 'dropout', the current study provides new insights regarding the magnitude of the return rate, and those students' immediate and longitudinal outcomes upon returning to college studies.

Contributions of the Study

This study contributes to the literature by showing that peer association can be evaluated outside of the traditional residential settings found at four-year institutions by using course enrollment as a proxy for student peer association. Evaluating peer association influence in settings different from four year, selective and residential colleges was considered a "fruitful area of future research" by Sacerdote (2001). This study has allowed for a measurement of student peer association at the curricular level using not only the number of peer associations, but also the rate/intensity of student peer-encounters, and determining the proportion of classmates which are peers. The methods presented in this study also allow an evaluation of the length of enrollment periods in which the student is either a soloist or attending with identified peers.

This study contributes to the literature by overcoming a significant internal constraint inherent to previous commuter college studies which used a randomized trial approach to measure the influence of student peer associations. As noted, randomized trials may provide a good estimation of the isolated peer effect, but are less applicable to the commuter and two-year college settings where students have limited interaction time while learning, or in activities outside the classroom setting (Center for Community College Student Engagement, 2012; Tinto & Russo, 1994; Townsend & Wilson, 2008; Visher, Schneider, Wathington, & Collado, 2010). In the settings of commuter institutions, students' entry level academic preparation is likely to be minimal but of equal level, and students will establish peer association in the natural settings without any assignment of certain peers. Randomized trials, such as the ones in the armed forces academies assign students to be as 'far apart' as possible: relatives, students from the same high school or town, and even students with the same last name would not be assigned to the same group. In the natural settings of a community college, a student may select their peers and those will defer from a random assignment in terms of their academic and demographic profile.

The evaluation of student-level transactional data found in administrative student records systems allows for the identification of various student behaviors, and is immune to the effect of student's non-response selection bias (Porter et al. (2004); (Massey & Tourangeau, 2013a, 2013b)). Longitudinal transaction data adds to previous studies' in its ability to measure the timing of registration behavior, and verify students' coincident enrollment in courses. In addition, the study presented a peer association measurement that fits well with the setting of commuter and two-year community colleges, where students have limited interaction time in learning activities outside the classroom setting (Center for Community College Student Engagement, 2012; V. Tinto & Russo, 1994; Townsend & Wilson, 2008; Visher et al., 2010), and often do not participate in the limited socializing opportunities provided through extra-curricular activities.

This study further contributes to the literature by demonstrating that measurements of peer association and of student-level behaviors can be derived from existing, transactional datasets. Peer association can take place using a retrospective dataset and allows the analysis to be conducted, and reduces the time constraint involved in a prospective randomized trials. Randomized, prospective trials may be limited by the sample size, college setting, and length of time between its initiation and the availability of the findings. The generation of these indicators does not require the administration of additional student surveys or other intrusive collection methods: Neither does it cost very much time or money to perform. This study demonstrates the benefit of linking individual student-level data drawn from a wide range of reliable and concurrent sources (entrance exams, transactional data, college transcripts, academic calendar information, course schedule and financial aid data). The ability to securely link multiple records from the academic, financial and administrative sections of colleges allows building an accurate view of students' attachment to the college, peer association network and progress towards his/her academic milestones.

Previous studies have employed late registration definitions that were not comparable across institutions and examined limited groups of students or over a limited period of time. This study developed what might be a more useful and perhaps universally comparable definition of late registration, and documented the magnitude of its occurrence not only at college entry, but also across time. Using the student transactional data, this study documented that nearly one in four students (718 out of 3,218, or 23%) at a large, urban, commuter community college, practiced late registration at some point during their studies. In addition to supporting the findings of previous studies regarding late registration's negative effect on GPA (i.e. Ford et al., 2008; Neighbors, 1996; Safer, 2009; Smith et al., 2002; Summers, 2000) and lower retention and completion rates (i.e. Summers, 2000, Smith et. al 2002, Freer‐Weiss: 2004, Johnson: 2006, Hiller 2005), this study found late registration was correlated with both immediate and long-term negative consequences to student academic outcomes. Previous studies found negative influence of late registration mostly for the immediate semester, as they rarely evaluated the longitudinal influence of late registration as an indicator. This study found that late registration is associated with students' negative outcomes beyond the semester in which the lateness took place. Furthermore, this study found a correlation between late registration and the likelihood of a student being either a soloist, and/or having few peer associations compared to the general population of the study cohort. There was a direct link between students' late registration behavior and their ability to influence academic outcomes through peer associations.

This study provides enhanced controls for the influence of student demographic, academic and financial indicators through a reevaluation by updating those variables, and re-running the regressions each semester. While other studies rely mainly the measurement of students' qualities at baseline, a revisit of student qualities each semester ensures a higher level accuracy. Many covariates related to student demographics are likely to remain fixed (e.g. gender or race/ethnic self-identification), while others, such as age, will change constantly; or such as student visa status, will change frequently. Over the course of the follow up period, nearly 22% of the students on international student visa have changed to or from a different status. These changes may seem minor, but the examples given to illustrate the consequences of changes that are tied to a student's eligibility for financial aid; particularly associated with need-based eligibility and dependency status. In the case of student visa status, changes are likely to reflect on a student's commitment to their studies, or the ranking of such commitment in relation to other competing commitments the student would face. To retain their visa, a student holding a student-visa status is required to maintain a minimal course load enrollment (typically, full time equivalency) and maintain a good academic standing. A resident student, or student on other visas is not committed to their academic coursework as a contingent condition to their permission to enter or remain in residence. Changing a visa status from student visa to other categories suggests that the student commitment to their studies may vary, as it is no longer a contingent requirement to their status. Furthermore, this study contributes to the literature and shows that non-resident student visa is not a 'stable' category that can be accurately captured by using entering semester status exclusively: nearly 25 percent of student visa holders have changed in or out of a student visa category during the study period.

This study addresses a growing concern regarding student response rates to survey and non-response bias. In a series of studies, Adams & Umbach (2012), Sax, Gilmartin, & Bryant (2003), Porter et al (2004, 2005, 2006) and Massey & Tourangeau, (2013a, 2013b) found survey completions rate are declining, and that the quality of the responses collected (completed or partial) is subject to a bias as well. This study provides new insights regarding the replacement of survey collections with existing information sources available to colleges: Using student transactional data allows for a near complete collection of indicators for student behavioral data that reflects directly on their actions and outcomes at the college. This collection can take place without adding any survey instruments, and provides the ability to measure relevant student-level indicators in a near complete coverage rate, without being subject to the student's or respondent interpretation of the question (e.g. would changes that are conducted due to class cancellation and can further replace or complete existing data collections.

Implications

The study carries implications to college administrators and faculty, institutions and for policy makers. The study shows that establishing and maintaining even a small group of peers' is associated with increased academic success. The peer association network benefits both from the number of peer associations (the number of peers) and from the intensity (repeated encounters with unique peers). An increase in the number and intensity of peer associations influences student's success. There may also be some form of value-added skills development related to planning and organization a student may acquire during their studies. Students who complete their planned path, register on time and retain their schedule are more likely to persist and have better academic outcomes.

College administrators, faculty and institutions should consider means that encourage student peer associations in addition to learning communities. While courses full of students in learning communities were found to have a positive influence on student outcomes, even the smaller networks of peers that organically emerged also resulted in positive outcomes. Furthermore, smaller networks are easier to maintain and support compared with the complex structure and scheduling of learning communities. The settings of learning communities presents complicated block scheduling, room allocations, advisors and space constraints on an institution, and do not fit the dynamic schedule constraints (and schedule rotations) a student may face. Supporting small networks of students with simpler means may provide a similar effect to student success. For example, it would be simpler for community colleges to generate a proposed schedule for students with the same 4-5 peers, than it would be to generate a proposed 'block' schedule for a group at a size of a typical classroom (25 to 30 students).

While several college systems and policymakers have suggested that institutions should eliminate late registration (i.e. Ignash, 1997; Boylan, Bonham & White, 1999; Roueche and Roueche, 1999; Lucy‐Allen, Merisotis, & Redmond, 2002; McClenney ,2004) this study does not take a position with regards to the policy (Specifically, this study does not imply that late registration should necessary be banned). Instead, the study suggests that late registration patterns should be taken into account once student risk-indicators (or risk profiles) are assessed. Further, this study demonstrates that student behavioral, academic, and financial profile is inevitably changing throughout their enrollment. For colleges that do assess students at entry, they might consider revaluating the student risk composite for each semester or academic year. This would allow community colleges to concentrate the limited resources available for student support services and advisement on the neediest student groups.

This study highlights that community colleges might benefit from monitoring student registration patterns closely. In only two of the studies reviewed (Broadbent, 1975; Hagedorn et al., 2007), were course schedule changes as a student behavior considered. This study documents the correlation of student-initiated schedule changes with student academic outcomes. It shows that schedule changes patterns can be examined and colleges may determine their applicability to the student's risk profile. In terms of data collection, the study informs decision makers of the wealth of information available through existing transactional data systems and its ability to address gaps in the understanding of student peer associations and other dynamic patterns while enrolled. These records can be mined for crucial data which may replace/augment the information collected through student surveys. When student-based surveys are employed, administrative records are useful in evaluating the accuracy of self-reported responses by comparing measures of student intentions with those of actual student behavior and academic outcomes. The mismatch between student reported responses and actual performance is relatively large and well documented. For example, NCES conducted a matching process of all Beginning Postsecondary Students (BPS2004/09) Studies with the National Student Clearinghouse and found the percent of discrepancies between student self-reporting and actual student actions reached up to 35% depending on the variable being examined (NCES: 2008, Wine, Janson, and Wheeless, 2011). This discrepancy rate suggests that transactional records can provide more accurate measurement of student behaviors compared with the student self-reported responses. While transactional data cannot replace entirely statewide or national data collections, colleges may use information from the transactional file to compile information regarding the students' behavior on a real time basis

Naturally, national statistical agencies, such as NCES, cannot assess in full the quality of administrative/transcript data provided as it is prepared outside the agency (Groen, 2012). But, colleges and universities can increase the data quality and consistency through the course of the student enrollment using the transactional data. Transactional files allow to examine the source of each change, the time of its occurrence and the proper coding of any change in the student record, resulting in higher completeness level and higher accuracy level compared with survey collection.

Furthermore, the study informs colleges' administrators, faculty and decision makers of the need to reexamine students' at-risk classification patterns not only on the basis of their pre-college or first semester data, but also across time. Using data as it evolves over time will allow colleges to better align the available resources for student support with students in need.

Limitations

This study may be limited by concerns regarding selection on unobservable variables. This limitation is similar to previous studies that estimated peer association influence in a non-randomized setting. As noted, this concern is centered on the selection of courses (type) which may differentiate the student's (and peers) abilities. (Arcidiacono et al., 2012). This study is not as likely to be influenced by this limitation, for two reasons. First, multiple offerings of foundation and remedial courses are offered in multiple course sections. Students were unlikely to self-select (as a group) to concentrate in a single course section. Even if students would concentrate at high-demand times during the week (e.g. mid-morning, midweek classes), they are presented with multiple, concurrent sections of the same foundation courses. Second, the models took into account a rich set of longitudinal, academic, demographic and financial covariates assembled from various reliable sources. Furthermore, the covariates were constantly updated to account for students' qualities prior to the first day of each semester. It allowed this study to account for numerous covariates unavailable to other studies due to data collection, specification, granularity, and survey size limitations.

Second, while the analyses provided a detailed picture of the links between the number (and intensity) of peer associations and the influence on student outcomes, it cannot establish the direction of the relationship. For example, although the findings show the increased number of peers is associated with increased credits earned, it is possible that students who have more peers simply attempted more courses (and credits). This bias potentially increases the influence of peer association due to a *correlation effect* (Manski (1993), Sacerdote (2011)). Despite this limitation, the study contributes to the understanding of peer association, as it controls for student covariates *prior* to changes in peer association. In addition, given the settings of a large, urban, community college, students are likely to have multiple class section offerings to select from when considering identical course taking requirements. Yet, the association between peers and academic outcomes remains significant in the presence of controls for student academics.

Third, this study was not concerned with the attributes related to the academic preparation of student peers. Previous studies, conducted at selective or highly selective colleges, have suggested that students' peer-association influence will vary by the quality of the peers. The setting of this study does not remove the influence of peer qualities variation. However, the natural setting at the open admission community college where this study took place enrolls a student body that minimizes the variation - nearly 85.0% of entering students required remediation in at least one academic area, and of those students holding a high school diploma 75% had an average of 80 or lower in high school. This limitation may be addressed in future studies.

Fourth, this study did not record a student reasoning for late registration or schedule changes. This study was better suited to measure schedule changes, as the transactional data enabled separation of student initiated changes from campus initiated changes. At the same time, the record does not indicate a reason given by the student for initiating the change. A better understanding of the motivation to initiate schedule changes may allow them to be better associated with the behavioral indicators directly, or with other factors, e.g. financial constraints. Though the study did not collect information regarding the reason for which the student imitated the change, it controls for a detailed, rich set of concurrent covariates covering multiple aspects of student's academics, demographics and financial indicators. Once those covariates were held constant, the influence of schedule changes and late registration remained in place, suggesting that the influence of those behaviors occurs even without a direct measurement of the reason leading the student to register late or change their schedule.

Future Research

This study demonstrated that comprehensive transactional records can provide a wealth of information regarding students' progress in community colleges. The study further shows that student peer association levels may be properly assembled and evaluated, resulting in fruitful indicators of students association and its influence on future outcomes. Future studies may benefit from evaluating the influence of class schedule during the week (daytime, evening or mixed schedule of classes) on student's peer association. It is possible that students who concentrate their coursework in daytime classes work fewer hours, or may be full time students and, during their studies, be out of the labor force (OLF).

While peer association was related to academic quality in each semester it will not necessary vary to an extreme, future studies may benefit from testing the influence of peer association quality. Given the settings of open admission, commuter, community colleges, studies should consider measures of academic quality that are not dependent on high school or pre-college student records (which may not reflect on student preparation if they are have delayed their high school enrollment). Other measures (e.g. credits earned to date or GPA) may reflect better on peers' academic qualities.

Third, the influence of cumulative peer association may be further evaluated through an examination of the student's peer association history: recent peers (e.g. current or previous semester) may have greater influence on student outcomes compared with long forgotten historical peers (those with whom the target student has had no recent encounters). As students move in and out of peer relationship they carry around them something resembling a "peer cloud" of associations. From a historical or longitudinal perspective a peer cloud represents those peers a target student has surrounded themselves with over several semesters. The overlap of being in the same courses with a consistent set of peers over a protracted period of time may yield new insights into the correlation of being at the center of this type of social circles and individual academic outcomes. The hypothesis here would be that larger peer clouds indicate higher peer association intensity compared to the lack of a peer cloud. Tracking these peer clouds, and the potential for the intersection of peer clouds between different individuals, would be likely provide new insights and fruitful information regarding students' interaction and their academic progress. Similarly, the relationship of individual qualities among students in peer-associations holds tremendous promise for further investigation.

Fourth, future studies should considered utilizing the exponential growth of digital information to evaluate students' interaction (on top of course enrollment information). For example, Marmaros and Sacerdote (2006); Mayer and Puller (2008) and Sacerdote (2011) have used email exchange data and social media data and suggested that this information may be used to evaluate both peer associations and patterns in the association (e.g. by gender/ethnicity, class year, proximity, etc.).

Lastly, future studies may benefit from a linkage of the student record to labor market outcomes. Future studies may achieve two major goals using this information. First- it would enable evaluating the student's workload during their studies (as reported on administrative records, e.g. financial aid application or tax returns). Secondly, it would allow the evaluation of students' future job market outcomes by peer association levels, and in relation to other behavioral indicators. As practices of linking student-level records to employment and/or tax records evolves, future studies would be able to evaluate the influence of student employment during their studies. Future studies would also be able to evaluate the influence of peer associations, and other behavioral indicators on labor market outcomes (placement rates and unemployment, and future income).

References.

Adams, M. J., & Umbach, P. D. (2012). Nonresponse and Online Student Evaluations of Teaching: Understanding the Influence of Salience, Fatigue, and Academic Environments. *Research in Higher Education, 53*(5), 576-591. doi:10.1007/s11162-011-9240-5

Adelman, C. (2005). Moving into Town--and Moving On: The Community College in the Lives of Traditional-Age Students. *US Department of Education*, 202.

Albright, J. J., & Marinova, D. M. (2010). Estimating multilevel models using SPSS, Stata, SAS, and R.

Angelo, D. T. (1990). The relationship between late registration and student persistence and achievement. *College and University, 64*, 316-327.

Angrist, J., Lang, D., & Oreopoulos, P. (2009). Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, 136-163.

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics,* *http://dx.doi.org/10.1016/j.labeco.2014.05.008*.

Arcidiacono, P., Foster, G., Goodpaster, N., & Kinsler, J. (2012). Estimating spillovers using panel data, with an application to the classroom. *Quantitative Economics, 3*(3), 421-470.

Baepler, P., & Murdoch, C. J. (2010). Academic analytics and data mining in higher education. *International Journal for the Scholarship of Teaching and Learning, 4*(2), 17.

Bahr, P. R. (2010). The bird’s eye view of community colleges: A behavioral typology of first-time students based on cluster analytic classification. *Research in Higher Education, 51*(8), 724-749.

Baker, R. S., Pechenizkiy, M., Romero, C., & Ventura, S. (2010). *Handbook of educational data mining*: CRC Press.

Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *JEDM-Journal of Educational Data Mining, 1*(1), 3-17.

Bean, J. P., & Metzner, B. S. (1985). A Conceptual Model of Nontraditional Undergraduate Student Attrition. *Review of educational research, 55*(4), 485-540.

Belcher, M. J., & Patterson, C. (1990). *Who Are Late Registrants and What Will They Do when Faced with a Late Registration Fee?* Miami, FL: Miami-Dade Community College. (ERIC Document Reproduction Service No. ED 328324).

Berger, J. B., & Milem, J. F. (2000). Organizational behavior in higher education and student outcomes. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 15, pp. 268-338). New York: Agathon.

Berland, M., Baker, R. S., & Blikstein, P. (2014). Educational data mining and learning analytics: Applications to constructionist research. *Technology, Knowledge and Learning, 19*(1-2), 205-220.

Broadbent, W. A. (1975). *An Inquiry into the Factors Underlying Class Schedule Changes*. Retrieved from Pearl City, HI:

Cabrera, A. F., Nora, A., & Castaneda, M. B. (1993). College Persistence: Structural Equations Modeling Test of an Integrated Model of Student Retention. *The Journal of Higher Education, 64*(2), 123-139.

Calderwood, L., & Lessof, C. (2009). Enhancing longitudinal surveys by linking to administrative data. *Methodology of longitudinal surveys*, 55-72.

Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *EDUCAUSE review, 42*(4), 40.

Carrell, S. E., Fullerton, R. L., & West, J. E. (2009). Does your cohort matter? Measuring peer effects in college achievement. *Journal of Labor Economics, 27*(3), 439-464.

Carrell, S. E., Sacerdote, B. I., & West, J. E. (2013). From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica, 81*(3), 855-882. doi:10.3982/ECTA10168

Center for Community College Student Engagement. (2012). *A Matter of Degrees: Promising Practices for Community College Student Success--A First Look*. Austin, TX. Retrieved from http://files.eric.ed.gov/fulltext/ED529070.pdf.

Chesebro, J. W., Snider, K., Venable, A., Green, R., & Mino, M. (1999, June). *Measuring Learning Community Effectiveness: Conceptions, an Instrument, and Results*. Paper presented at the Association of Higher Education (AAHE), Denver, CO.

Chilton, B. S. (1964). *The relationship between certain factors of late and regular registrants.* North Texas State University.

Conklin, K. A. (1997). Course attrition: A 5-year perspective on why students drop classes. *Community College Journal of Research & Practice, 21*(8), 753.

Cook, T. D., Campbell, D. T., & Shadish, W. (2006). *Experimental and quasi-experimental designs for generalized causal inference*: Houghton Mifflin.

Cornille, K. T. (2009). *The Impact Of Late Registration On Academic Success And Persistence Of Students At A Community And Technical College.* (Ed.D.), Edgewood College, Ann Arbor.

Crosnoe, R., Riegle-Crumb, C., Field, S., Frank, K., & Muller, C. (2008). Peer group contexts of girls' and boys' academic experiences. *Child development, 79*(1), 139-155.

Czajka, J. L. (2013). Can Administrative Records Be Used to Reduce Nonresponse Bias? *The Annals of the American Academy of Political and Social Science, 645*(1), 171-184. doi:10.1177/0002716212463313

Desimone, J. (2007). Fraternity membership and binge drinking. *Journal of Health Economics 26*, 950-967.

DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (1999). An event history model of student departure. *Economics of Education Review, 18*(3), 375-390. doi:Doi: 10.1016/s0272-7757(98)00049-1

Dey, E. L. (1997). Working with Low Survey Response Rates: The Efficacy of Weighting Adjustments. *Research in Higher Education, 38*(2), 215-227. doi:10.1023/A:1024985704202

Diablo Valley College. (2004). Impact of Student Registration Behavior on Academic Performance.   Retrieved from http://www.dvc.edu/org/departments/research/pdf/Impact%20of%20Student%20Registration%20Behavior.pdf

Diekhoff, G. M. (1992). Late registrants: At risk in introductory psychology. *Teaching of Psychology, 19*(1), 49-50.

Duflo, E., Pascaline, D., & Kremer, M. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review, 101*(5), 1739-1774. doi:10.1257/aer.101.5.1739

Duncan, G. J., Boisjoly, J., Kremer, M., Levey, D. M., & Eccles, J. (2005). Peer Effects in Drug Use and Sex Among College Students. *Journal of Abnormal Child Psychology, 33*(3), 11. doi:10.1007/s10802-005-3576-2

Engstrom, C., & Tinto, V. (2008) Learning Better Together: The Impact of Learning Communities on the Persistence of Low-Income Students. *Vol. 1. Opportunity Matters* (pp. 5-21). Washington, DC: Pell Institute for the Study of Opportunity in Higher Education. Retrieved from http://www.pellinstitute.org/journal/Pell_OppMattersv1_EngstromTinto.pdf.

Feld, J., & Zölitz, U. (2014, June 2014). On the Nature of Peer Effects in AcademicAchievement. *Working Papers in Economics.* June 2014.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning, 4*(5), 304-317.

Flemming, E. (1992). *NCES Statistical Standards*. Washington, DC National Center for Education Statistics.

Regulations of the University of Florida, § 4 (2012).

Ford, G. G., Stahl, K. J., Walker, M. E., & Ford, A. M. (2008). Better late than never? The relation of registration date to class performance. 42: 402-407. *College Student Journal, 42*, 402-407.

Foster, G. (2006). It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism. *Journal of public Economics, 90*(8), 1455-1475.

Freer-Weiss, D. (2004). Community college freshmen: Last in, first out? *Journal of College Student Retention: Research, Theory and Practice, 6*(2), 137-154.

Garlick, R. (2013). *Academic Peer Effects with Di referent Group Assignment Rules: Residential Tracking versusRandom Assignment*. Retrieved from Development Research Group at The World Bank and Duke University Departmentof Economics.:

Goodman, P. (2010). *Predictors of Persistence for First-Time, Full-Time Community and Technical College Students.* (Ph.D.), Walden University.

Griffith, A. L., & Rask, K. N. (2014). Peer Effects in Higher Education: A look at heterogeneous impacts. *Economics of Education Review. Retrieved from* *http://ac.els-cdn.com/S0272775714000168/1-s2.0-S0272775714000168-main.pdf?_tid=f5f6d6dc-b213-11e3-9d5d-00000aacb35f&acdnat=1395528635_6c38740b3a636464860f147fcd8b9470*.

Grimes, S. K., & Antworth, T. (1996). Community college withdrawal decisions: Student characteristics and subsequent reenrollment patterns. *Community College Journal of Research and Practice, 20*(4), 345-361.

Groen, J. A. (2012). Sources of error in survey and administrative data: The importance of reporting procedures. *Journal of Official Statistics, 28*(2), 173.

Hagedorn, L. S., Maxwell, W. E., Cypers, S., Moon, H. S., & Lester, J. (2007). Course shopping in urban community colleges: An analysis of student drop and add activities. *The Journal of Higher Education, 78*(4), 464-485.

Hale, J. M., & Bray, N. J. (2011). The impact of registration timing on student performance. *Community College Journal of Research and Practice, 35*(7), 556-573. doi:10.1080/10668920802289984

Hamilton, L. (2012). *Statistics with Stata: version 12*: Cengage Learning.

Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does Peer Ability Affect Student Achievement? *Journal of Applied Econometrics, 18*, 527-544. doi:10.1002/jae.741

Heiner, C., Heffernan, N., & Barnes, T. (2007). *Educational data mining.* Paper presented at the Supplementary Proceedings of the 12th International Conference of Artificial Intelligence in Education.

Hiller, D. (2005). Late registrants: They are prepared and they do come back if successful. Retrieved from http://www.nmc.edu/ir/intell/analysis/lateregistrantssucces.pdf

Horn, L. (2009). *On Track to Complete? A Taxonomy of Beginning Community College Students and Their Outcomes 3 Years After Enrolling: 2003–04 Through 2006*. (NCES 2009-152). Washington, DC: National Center for Education Statistics, Institute for Education Sciences, U.S. Department of Education. Retrieved from http://nces.ed.gov/pubs2009/2009152.pdf.

Horn, L. (2009). On track to complete?: a taxonomy of beginning community college students and their outcomes 3 years after enrolling: 2003-04 through 2006.

Horn, L., & Nevill, S. (2006). *Profile of Undergraduates in U.S. Postsecondary Education Institutions: 2003–04: With a Special Analysis of Community College Students*. (NCES 2006-184). U.S. Department of Education, Washington, DC: NCES Retrieved from National Center for Education Statistics http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2006184.

Hu, S., & McCormick, A. (2012). An Engagement-Based Student Typology and Its Relationship to College Outcomes. *Research in Higher Education, 53*(7), 738-754. doi:10.1007/s11162-012-9254-7

Johnston, G. H. (2006). *The Date of Course Enrollment as a Predictor of Success and Persistence*. Paper presented at the 46th Annual Forum of the Association for Institutional Research (AIR), Chicago, IL, May 14-18, 2006. www.eric.ed.gov/ERICWebPortal/recordDetail?accno=ED493834

Junco, R., & Cotten, S. R. (2011). Perceived academic effects of instant messaging use. *Computers & Education, 56*(2), 370-378. doi:http://dx.doi.org/10.1016/j.compedu.2010.08.020

Karp, M. M. (2011). Toward a New Understanding of Non-Academic Student Support: Four Mechanisms Encouraging Positive Student Outcomes in the Community College. CCRC Working Paper No. 28. Assessment of Evidence Series. *Community College Research Center, Columbia University*.

Keck, K. L. (2007). *Community College Retention: The Role of Late Registration Policies.* (Ph.D.), Western Michigan University, Kalamazoo, Michigan.

Knauf, R., Sakurai, Y., Takada, K., & Tsuruta, S. (2010). *Validation of a Data Mining Method for Optimal University Curricula.* Paper presented at the International Workshop on Design, Evaluation and Refinement of Intelligent Systems (DERIS2010).

Knight, W. E. (2003). Learning Communities and First-Year Programs: Lessons for Planners. *Planning for Higher Education, 31*(4), 5-12.

Koenig, M. L. (2003). *An assessment of the Current Population Survey and the Survey of Income and Program Participation using Social Security administrative data.* Paper presented at the Federal Committee on Statistical Methodology Research Conference, Washington, DC, November.

Luan, J. (2002). Data mining and its applications in higher education. *New Directions for Institutional Research, 2002*(113), 17-36.

Luan, J. (2007). Data Mining Applications in Higher Education: Executive Report SPSS, http://www. pse. pt/Documentos/Data% 20mining% 20in% 20higher% 20education. pdf.

Lyle, D. S. (2007). Estimating And Interpreting Peer And Role Model Effects From Randomly Assigned Social Groups At West Point. *The Review of Economics and Statistics, 89*(2), 11.

Mannan, G., & Preusz, G. C. (1976). University Late Registrants and Related Academic Problems. *College Student Journal*.

Manski, C. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies, 60*(3), 11.

Marmaros, D., & Sacerdote, B. (2006). How Do Friendships Form? *The Quarterly Journal of Economics, 121*(1), 79-119.

Massey, D. S., & Tourangeau, R. (2013a). Introduction: New Challenges to Social Measurement. *The Annals of the American Academy of Political and Social Science, 645*(1), 6-22. doi:10.1177/0002716212463314

Massey, D. S., & Tourangeau, R. (2013b). Where Do We Go from Here? Nonresponse and Social Measurement. *The Annals of the American Academy of Political and Social Science, 645*(1), 222-236. doi:10.1177/0002716212464191

Mayer, A., & Puller, S. L. (2008). The old boy (and girl) network: Social network formation on university campuses. *Journal of public Economics, 92*(1), 329-347.

McWaine, W. L. (2012). Late Registration and African American Males' Academic Performance in a Suburban Community College System. *ProQuest LLC*.

Mendiola-Perez, C. L. (2004). *The effects of early, regular, and late registration on academic success and retention of community college students.* Baylor University.

Moore, C., Shulock, N., Ceja, M., & Lang, D. M. (2007). *Beyond the Open Door: Increasing Student Success in the California Community Colleges*. Retrieved from Sacramento, CA:

Moran, J. J., Bausili, M. T., & Kramer, M. (1995). A Look at the Reasons Behind the Drop/Add Process. *College and University, 71*(1), 2-5, 7-11.

Morris, E. W. (1986). A Study of Changes of Registration by Adding and Dropping Classes. *College and University, 61*(4), 327-337.

NCES. (2012a). *2004/2009 Beginning Postsecondary Students Longitudinal Study Restricted-Use Transcript Data Files and Documentation*.

NCES. (2012b). *NCES Statistical Standards* (3 ed.). Washington, DC: NCES.

Neighbors, J. E. G. (1996). *The impact of early. regular, and late registration on students at three Texas colleges.* (Ed.D. Unpublished Dissertation), Texas A & M University-Commerce, Commerce, TX.

Newcomb, T. M. (1962). Student peer-group influence. In J. Adelson (Ed.), *The American college* (pp. 469-488). New York: John Wiley.

Newcomb, T. M., Brown, D. R., Kulik, J. A., Reimer, D. J., & Revelle, W. R. (1970). Self-selection and change. In J. Gaff & Associates (Eds.), *The cluster college* (pp. 137-160). San Francisco:: Jossey-Bass.

O'Banion, T. (2012). Late Registration: May It Rest in Peace. *Community College Journal, 83*(1), 26-31.

Parks, K. M. (1974). *An investigation of selected variables in regard to regular and late college registration.*

Pascarella, E. T., & Terenzini, P. T. (2005). *How College Affects Students : A Third Decade of Research* (2 ed. Vol. 2). San Francisco: Jossey-Bass.

Peña-Ayala, A. (2013). *Intelligent and Adaptive Educational-Learning Systems: Achievements and Trends*. New York Dordrecht London: Springer.

Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. *Expert systems with applications, 41*(4), 1432-1462.

Perkins, M. E. (2002). *Late Registration Policy and Implications for Student Success at a Single Community College.* (Ed.D.), University of Illinois at Urbana-Champaign, Urbana, IL.

Peterson, B. B. (1986). Institutional Research: A Follow-Up Study on the Late Applicant and the Late Registrant at Honolulu Community College. A Leaders for the 80's Project.

Peytchev, A. (2013). Consequences of Survey Nonresponse. *The Annals of the American Academy of Political and Social Science, 645*(1), 88-111. doi:10.1177/0002716212461748

Porter, S. R., & Umbach, P. D. (2006). Student survey response rates across institutions: Why do they vary? *Research in Higher Education, 47*(2), 229-247.

Porter, S. R., & Whitcomb, M. E. (2005). Non-response in student surveys: The Role of Demographics, Engagement and Personality. *Research in Higher Education, 46*(2), 127-152. doi:10.1007/s11162-004-1597-2

Porter, S. R., Whitcomb, M. E., & Weitzer, W. H. (2004). Multiple surveys of students and survey fatigue. *New Directions for Institutional Research, 2004*(121), 63-73. doi:10.1002/ir.101

Price, D. V. (2005). *Learning communities and student success in postsecondary education. A background paper*. MDRC. Retrieved from http://files.eric.ed.gov/fulltext/ED489439.pdf

Richburg-Hayes, L., Visher, M. G., & Bloom, D. (2008). Do Learning Communities Effect Academic Outcomes? Evidence From an Experiment in a Community College. *Journal of Research on Educational Effectiveness, 1*(1), 33-65. doi:10.1080/19345740701692472

Riegle-Crumb, C., Farkas, G., & Muller, C. (2006). The role of gender and friendship in advanced course taking. *Sociology of Education, 79*(3), 206-228.

Robins, J. M., Rotnitzky, A., & Zhao, L. P. (1994). Estimation of Regression Coefficients When Some Regressors Are Not Always Observed. *Journal of the American Statistical Association, 89*(427), 846-866. doi:10.2307/2290910

Rocconi, L. M. (2011). The impact of learning communities on first year students' growth and development in college. *Research in Higher Education, 52*(2), 178-193.

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert systems with applications, 33*(1), 135-146.

Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 40*(6), 601-618.

Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 3*(1), 12-27.

Roueche, J., & Roueche, S. (1993). *Between a Rock and a Hard Place: The At-Risk Student in the Open-Door College*. Washington, DC: American Association of Community Colleges.

Sacerdote, B. (2001). Peer Effects with Random Assignment: Results from Dartmouth Roommates. *The Quarterly Journal of Economics,* 25.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? *Handbook of the Economics of Education, 3*, 249-277. doi:10.1016/S0169-7218(11)03004-8

Safer, A. M. (2009). The Effect of Late Registration for College Classes. *College Student Journal, 43*(4), 9.

Sarraf, S., & Tukibayeva, M. (2014). Survey Page Length and Progress Indicators: What Are Their Relationships to Item Nonresponse? *New Directions for Institutional Research, 2014*(161), 83-97. doi:10.1002/ir.20069

Sax, L. J., Gilmartin, S. K., & Bryant, A. N. (2003). Assessing Response Rates and Nonresponse Bias in Web and Paper Surveys. *Research in Higher Education, 44*(4), 409-432. doi:10.1023/A:1024232915870

Scheuer, O., & McLaren, B. M. (2012). Educational data mining *Encyclopedia of the Sciences of Learning* (pp. 1075-1079): Springer.

Schmidt, M. M. (2004). *An examination of variables associated with student completion at Clovis Community College.* New Mexico State University.

Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating causal effects using experimental and observational designs*. Washington, DC: American Educational Research Association. Retrieved from http://www.aera.net/uploadedFiles/Publications/Books/Estimating_Causal_Effects/Causal_Effects.pdf.

Schuck, P. H., & Zeckhauser, R. (2006). *Targeting in social programs: Avoiding bad bets, removing bad apples*: Brookings Institution Press.

Schuh, J. H. (2011). *Assessment methods for student affairs*: John Wiley & Sons.

Scott, M. A., & Kennedy, B. B. (2005). Pitfalls in pathways: Some perspectives on competing risks event history analysis in education research. *Journal of Educational and Behavioral Statistics, 30*(4), 413-442.

Scrivener, S., Bloom, D., LeBlanc, A., Paxson, C., Rouse, C. E., & Sommo, C. (2008). Opening Doors. A Good Start: Two-Year Effects of a Freshmen Learning Community Program at Kingsborough Community College. *MDRC*.

Serban, A. M. (2002). Knowledge Management: The "Fifth Face" of Institutional Research. *New Directions for Institutional Research, 2002*(113), 105-112. doi:10.1002/ir.40

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin Company.

Shapiro, N. S., & Levine, J. H. (1999). *Creating Learning Communities: A Practical Guide to Winning Support, Organizing for Change, and Implementing Programs*. San Francisco, CA: Jossey-Bass.

Siemens, G., & Baker, R. S. (2012). *Learning analytics and educational data mining: towards communication and collaboration.* Paper presented at the Proceedings of the 2nd international conference on learning analytics and knowledge.

Sinclair Community College. (2005). *Paradigm Shift To Emphasize Student Success: Eliminating Late Registration After Classes Start*. Retrieved from http://www.sinclair.edu/about/administrative/vpss/pub/paradigm_shift.pdf.

Singer, J. D., & Willett, J. B. (1993). It's about Time: Using Discrete-Time Survival Analysis to Study Duration and the Timing of Events. *Journal of Educational Statistics, 18*(2), 155-195.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*: Oxford university press.

Smart, J. C. (Ed.) (2010). *Higher Education: Handbook of Theory and Research* (Vol. XXV): Springer Science + Business Media B.V. http://www.springerlink.com/content/j11g28/front-matter.pdf.

Smith, A. B., Street, M. A., & Olivarez, A. (2002). Early, regular, and late registration and community college student success: A case study. *Community College Journal of Research and Practice, 26*(3), 261-273. doi:10.1080/106689202317245455

Smith, J., & Stange, K. (2015). *A New Measure of College Quality to Study the Effects of College Sector and Peers on Degree Attainment*. Retrieved from

Sova, A. D. (1986). *A Study of the Success Rate of Late Admits in Freshman English at the Two-Year College.* (Working Paper 2-86). Binghamton, NY: Broome Community College, Institute for Community College Research.

Stassen, M. L. A. (2003). Impact of Varying Living-Learning Community Models. *Research in Higher Education, 44*(5).

StataCorp. (2014). *Stata Statistical Software: Release 14.* College Station, TX: StataCorp LP.

Stein, J. B. (1984). Should Late Enrolling New Students Be Registered for Classes? *Minneapolis, MN: Minneapolis Community College. (ERIC Document Reproduction Service No. ED 243537).*

Stinebrickner, R., & Stinebrickner, T. R. (2006). What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds. *Journal of public Economics, 90*(8), 1435-1454. doi:10.1016/j.jpubeco.2006.03.002

Street, M. A. (2000). *The Effects of Early, Regular, and Late Registration on Student Success in Community Colleges.* (Ed.D.), Texas Tech University, Lubbock, TX.

Summers, M. D. (2000a). *Enrollment and Registration Behaviors as Predictors of Academic Outcomes for Full-Time Students in a Rural Community College.* (Ed.D.), University of Illinois at Urbana-Champaign, Urbana, IL.

Summers, M. D. (2000b). *Enrollment and Registration Behaviors as Predictors of Academic Outcomes for Full-Time Students in a Rural Community College*. Paper presented at the 42nd Annual Conference of the Council for the Study of Community Colleges, Washington, D.C.

Summers, M. D. (2003). Attrition Research at Community Colleges. *Community College Review, 30*(4), 64.

College Finance and Business Operations (College Operating Fund) for the State University of New York, 602 C.F.R. (2015).

Taylor, K. (2003). *Learning community research and assessment: What we know now*: Washington Center for Improving the Quality of Undergraduate Education, National Learning Communities Project.

Thomas-Spiegel, J. (1997). *Increasing Instructor and Student Communication through Drop Notice Surveys*. Retrieved from http://files.eric.ed.gov/fulltext/ED409923.pdf

Tincher-Ladner, L. (2006). *A Review of Late Registration Participants of Mississippi Gulf Coast Community College*. Retrieved from Perkinston, MS: http://themediafilegroup.com/downloadpage/late-registration.pdf.html?wm=165&sub=1&u=IS6w4eTOHaCZjG1mrFee_hZJ2NnHvEmxGUwiYsq2tJOY.

Tinto, V. (1993). *Leaving College: Rethinking the Causes and Cures of Student Attrition* (2 ed.). Chicago, IL: The University of Chicago Press.

Tinto, V. (1997). Classrooms as Communities: Exploring the Educational Character of Student Persistence. *The Journal of Higher Education, 68*(6), 599-623. Retrieved from http://www.jstor.org/stable/2959965.

Tinto, V. (2003). Learning better together: The impact of learning communities on student success. *Higher Education monograph series, 1*(8).

Tinto, V., & Russo, P. (1994). Coordinated Studies Programs: Their Effect on Student Involvement at a Community College. *Community College Review, 22*(2), 16-25. doi:10.1177/009155219402200203. Retrieved from http://crw.sagepub.com/content/22/2/16.full.pdf+html

Townsend, B. K., & Wilson, K. B. (2008). The Academic and Social Integration of Persisting Community College Transfer Students. *Journal of College Student Retention, 10*(4), 405-423. doi:10.2190/CS.10.4.a

Tsuruta, S., Knauf, R., Dohi, S., Kawabe, T., & Sakurai, Y. (2013). An Intelligent System for Modeling and Supporting Academic Educational Processes. In A. Peña-Ayala (Ed.), *Intelligent and Adaptive Educational-Learning Systems* (Vol. 17, pp. 469-496): Springer Berlin Heidelberg.

U.S. Department of Education. (2009a). *Integrated Postsecondary Education Data System, Enrollment File*. Washington, DC: NCES.

U.S. Department of Education. (2009b). *Integrated Postsecondary Education Data System, Finanacial Aid File*. Washington, DC: NCES.

Visher, M. G., Schneider, E., Wathington, H., & Collado, H. (2010). Scaling up Learning Communities: The Experience of Six Community Colleges. *National Center for Postsecondary Research*, 153.

Visher, M. G., Wathington, H., Richburg-Hayes, L., & Schneider, E. (2008). The Learning Communities Demonstration: Rationale, Sites, and Research Design. An NCPR Working Paper. *National Center for Postsecondary Research*.

Volkwein, J. F. (1999). The Four Faces of Institutional Research. *New Directions for Institutional Research, 1999*(104), 9-19.

Weidman, J. (Ed.) (1989). *Undergraduate socialization: A conceptual approach* (Vol. 5). New York: Agathon.

Weiss, M. J., Mayer, A., Cullinan, D., Ratledge, A., Sommo, C., & Diamond, J. R. (2014). A Random Assignment Evaluation of Learning Communities at Kingsborough Community College – Seven Years Later. *Journal of Research on Educational Effectiveness*. doi:10.1080/19345747.2014.946634

West, D. M. (2012). Big data for education: Data mining, data analytics, and web dashboards. *Governance Studies at Brookings*, 1-10.

Wilson, J. (2007). Peer Effects and Cigarette Use Among College Students. *Atl Econ Journal, 35*, 15. doi:DOI 10.1007/s11293-007-9064-z

Winkler, W. E. (2001). Record linkage software and methods for merging administrative lists. *US Bureau of the Census*.

Winston, G., & Zimmerman, D. (2004). Peer effects in higher education *College choices: The economics of where to go, when to go, and how to pay for it* (pp. 395-424): University of Chicago Press.

Zhao, C.-M., & Kuh, G. D. (2004). Adding Value: Learning Communities and Student Engagement. *Research in Higher Education, 45*(2), 115-138.

Zimmerman, D. (2003). Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. *The Review of Economics and Statistics, 85*(1), 16.

Zottos, G. (2005). *The Impact of Late Registration on Academic Outcomes of Urban Community College Students.* (Ed.D.), University of Southern California, Los Angeles, CA.

APPENDIX A. Revisit the Correlation between Peer Association and Course Load in the First Semester

A review of table 15 suggests that the number of peer associations does not increase at the same rate of the number of courses. Perhaps some of the peer associations may be a result of learning communities' course bundles, and some peer associations may be a result of chance alone. Table A.1 employs a slightly stricter definition of peer association, where students need to meet *three times or more* to be considered as having peer association.

Table A.1

Proportion of Students and the Number of Peers Associations (meeting three times or more) by Number of Courses Taken in the First Semester.

	Courses Taken					
# of Peer Associations	**1**	**2**	**3**	**4**	**5+**	**Total**
0	161 (100.0%)	190 (100.0%)	295 (96.1%)	552 (83.5%)	1160 (64.2%)	2357 (75.4%)
1	0 (0.0%)	0 (0.0%)	8 (2.6%)	53 (8.0%)	155 (8.6%)	217 (6.9%)
2	0 (0.0%)	0 (0.0%)	1 (0.3%)	8 (1.2%)	41 (2.3%)	50 (1.6%)
3	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.2%)	7 (0.4%)	8 (0.3%)
4	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	6 (0.3%)	6 (0.2%)
5+	0 (0.0%)	0 (0.0%)	4 (1.3%)	47 (7.1%)	439 (24.3%)	490 (15.7%)
Total	161 (100.0%)	190 (100.0%)	308 (100%)	661 (100.0%)	1808 (100.0%)	3128 (100.0%)

An increase in the number of times peers meet to define an association from two to three increased the proportion of soloists from 18.8 to 75.4 percent. Not only the proportion of students with peer associations declines, but also the number of peers declines. The proportion of students with one peer declined by more than half from 14.2 to 6.9 percent. The proportion of students with five or more peers decline from 35.6 to 15.7 percent. The differences in the number and proportion of peers through a minor change in peer association definition suggests that some of the initial peer associations seen is a result of 'random' course 'bundles' resulting students taking classes together. For that reason, it is crucial to evaluate student peer association not only in the semester, but also across semesters.

APPENDIX B. Revisit the Number of Semesters Enrolled as a Soloist for Male Students

Table 27 reported the number of semesters a student enrolled during the study and the number of semesters enrolled as a soloist. The table reported the breakdown for female students. Table B.1. revisits the analysis, and reports the number of semesters a student enrolled during the study period and the number of semesters a student attended as a soloist by gender, for male students.

Table B.1.

Number of Semesters Enrolled In the Study Period, and Number of Semesters Enrolled As a Soloist (Male Students Only)

			Number of Semesters Enrolled as a Soloist									
		Total Males	0	1	2	3	4	5	6	7	8	9+
Total Semesters Enrolled	1	228	165	63								
	2	199	106	65	28							
	3	148	57	51	27	13						
	4	171	46	54	48	17	6					
	5	136	34	40	34	18	6	4				
	6	111	15	31	39	11	6	5	4			
	7	88	11	26	22	10	12	4	1	2		
	8	60	4	20	10	11	6	3	4	0	2	
	9+	60	0	2	9	11	8	10	10	5	4	1
		1201	438	356	217	91	44	26	19	7	6	1

As noted, the differences in soloist rates among female and male students was negligible. Two-thirds (63.5%) of the male students (n=763) were soloists for one or more semesters, compared with 64.3% for female students, and 64.2% for all students in the study cohort. Within the male population, 10.2% (n=123) were soloists throughout their enrollment, compared with 10.3% for female students, and 10.2% for the entire study cohort. Based on gender there were no significant differences in the number of semesters enrolled as a soloist.

Appendix C. A Revisit of the Correlation between Student Peer Association and Student Outcomes for Students with Minimal Courses Attempted

Students who took very few courses through their enrollment period may lack the opportunity to create student curricular peer association. The concern raised in such case is that the estimated coefficients for student peer association may be biased. To address this concern, Tables C.1. through Table C.6. revisit the models, limiting the sample only to students that took at least three courses throughout their studies. A total of n=241 students were removed for not meeting the minimal number of courses, leaving the sample with a total of 2,887 students instead of 3,128 students. Limiting the sample to students who took a minimal number of courses allows to further examine the correlation of student peer association as measured in the model.

The results demonstrate that the correlations reported for student peer association on students' academic outcomes (student semestrial GPA, semestrial credits earned, cumulative GPA, Cumulative credits earned and the odds of retention or graduation) remain consistent once all academic, demographic and financial covariates are taken into account. The correlations were not biased by students who did not have the opportunity to develop curricular peer associations. When the samples were limited to students who took at least three courses, a fractional change in the standard errors (within 0.01 or lower, e.g. 0.203 to 0.204) is reported. Those changes are attributed mainly to the reduction in the sample size.

Table C.1.

Students' Semestrial GPA by Peer Association, Additional Behaviors, Academic, Demographic, and Financial Indicators (Models 1 through 5, 5r-limited to students who took at least three courses)

Y=Semestrial GPA	Model 1	Model 2	Model 3	Model 4	Model 5	Model 5r
Behavioral Indicators						
Peer Associations (Encountered Twice)	0.033 *** (0.004)	0.032 *** (0.004)	0.013 *** (0.004)	0.012 ** (0.004)	0.012 ** (0.004)	0.011 ** (0.004)
Peer Associations (Encountered 3 times)	0.032 *** (0.006)	0.033 *** (0.006)	0.022 *** (0.006)	0.017 ** (0.006)	0.016 * (0.006)	0.015 * (0.006)
Peer Associations (Encountered 4+ times)	0.044 *** (0.005)	0.042 *** (0.005)	0.023 *** (0.005)	0.019 *** (0.005)	0.019 *** (0.005)	0.018 *** (0.005)
Late Registration		-0.216 *** (0.042)	-0.198 *** (0.042)	-0.177 *** (0.042)	-0.167 *** (0.042)	-0.142 *** (0.044)
Change Schedule		-0.176 *** (0.02)	-0.202 *** (0.021)	-0.203 *** (0.02)	-0.205 *** (0.02)	-0.193 *** (0.021)
Return from Stopout		0.145 *** (0.044)	0.250 *** (0.046)	0.243 *** (0.046)	0.248 *** (0.046)	0.260 *** (0.046)
Academic Indicators						
Need Writing (1st day of semester)			-0.134 *** (0.033)	-0.104 *** (0.033)	-0.097 ** (0.032)	-0.085 ** (0.033)
Need Math (1st day of semester)			-0.248 *** (0.036)	-0.192 *** (0.035)	-0.182 *** (0.035)	-0.194 *** (0.035)
Total Credits Earned (1st day of semester)			0.010 *** (0.001)	0.009 *** (0.001)	0.009 *** (0.001)	0.009 *** (0.001)
Credits attempted in the semester			-0.024 *** (0.004)	-0.020 *** (0.004)	-0.021 *** (0.004)	-0.021 *** (0.004)
Number of Classes in the semester			0.108 *** (0.012)	0.119 *** (0.012)	0.119 *** (0.012)	0.118 *** (0.013)
Part Time in the semester			-0.030 (0.034)	-0.028 (0.034)	-0.054 (0.034)	-0.033 (0.035)
Demographic Indicators						
Age				0.028 *** (0.002)	0.028 *** (0.002)	0.031 *** (0.002)
Male				-0.194 *** (0.035)	-0.196 *** (0.034)	-0.200 *** (0.035)
Race/Ethnic Profile (Reference=White)						
Int. Student				0.415 *** (0.053)	0.330 *** (0.057)	0.278 *** (0.058)
Asian				-0.050 (0.064)	-0.037 (0.063)	-0.076 (0.064)
Hispanic				-0.397 *** (0.055)	-0.362 *** (0.055)	-0.369 *** (0.055)
African Americans				-0.514 *** (0.062)	-0.479 *** (0.061)	-0.502 *** (0.062)
Other				-0.330 *** (0.059)	-0.312 *** (0.058)	-0.314 *** (0.059)
Financial Indicators						
Paid with Aid (Any Portion of Tuition)					0.430 *** (0.048)	0.399 *** (0.048)
Paid with Self (Any Portion of Tuition)					-0.031 (0.031)	-0.029 (0.031)
Percent Tuition Paid using Aid					-0.434 *** (0.062)	-0.382 *** (0.063)
Pell or Tap recipient					-0.170 *** (0.041)	-0.227 *** (0.042)
Time (Semesters)	-0.029 *** (0.005)	-0.035 *** (0.005)	-0.088 *** (0.008)	-0.087 *** (0.008)	-0.082 *** (0.008)	-0.089 *** (0.008)
Constant	2.163 *** (0.025)	2.266 *** (0.027)	2.225 *** (0.055)	1.776 *** (0.095)	1.866 *** (0.102)	1.871 *** (0.104)

Note: Coefficients (and Standard Errors) reported, *p≤0.05, **p≤0.01, ***p≤0.001

Table C.2.

Students' Credits Earned by Peer Association, Additional Behaviors, Academic, Demographic, and Financial Indicators (Models 1 through 5, 5r-limited to students who took at least three courses)

Y=Semestrial Credits Earned	Model 1	Model 2	Model 3	Model 4	Model 5	Model 5r
Behavioral Indicators						
Peer Associations (Encountered Twice)	0.673 *** (0.016)	0.647 *** (0.016)	0.079 *** (0.011)	0.077 *** (0.011)	0.077 *** (0.011)	0.075 *** (0.011)
Peer Associations (Encountered 3 times)	0.335 *** (0.028)	0.329 *** (0.028)	0.111 *** (0.018)	0.103 *** (0.018)	0.102 *** (0.018)	0.101 *** (0.018)
Peer Associations (Encountered 4+ times)	0.212 *** (0.021)	0.208 *** (0.021)	0.051 *** (0.015)	0.043 ** (0.015)	0.044 ** (0.015)	0.041 ** (0.015)
Late Registration		-2.319 *** (0.169)	-0.412 *** (0.106)	-0.351 *** (0.106)	-0.319 ** (0.107)	-0.315 ** (0.110)
Change Schedule		-0.561 *** (0.087)	-1.316 *** (0.055)	-1.307 *** (0.055)	-1.301 *** (0.055)	-1.310 *** (0.055)
Return from Stopout		-1.571 *** (0.176)	0.147 (0.111)	0.182 (0.111)	0.208 † (0.111)	0.213 † (0.112)
Academic Indicators						
Need Writing (1st day of semester)			-0.293 *** (0.084)	-0.251 ** (0.083)	-0.248 ** (0.083)	-0.255 ** (0.084)
Need Math (1st day of semester)			-0.215 * (0.088)	-0.103 (0.087)	-0.093 (0.087)	-0.104 (0.089)
Total Credits Earned (1st day of semester)			0.019 *** (0.002)	0.019 *** (0.002)	0.021 *** (0.002)	0.021 *** (0.003)
Credits attempted in the semester			0.746 *** (0.009)	0.749 *** (0.009)	0.745 *** (0.009)	0.742 *** (0.009)
Number of Classes in the semester			0.143 *** (0.033)	0.162 *** (0.033)	0.153 *** (0.033)	0.168 *** (0.033)
Part Time in the semester			-0.069 (0.091)	-0.070 (0.091)	-0.114 (0.092)	-0.091 (0.094)
Demographic Indicators						
Age				0.045 *** (0.006)	0.044 *** (0.006)	0.046 *** (0.006)
Male				-0.315 *** (0.083)	-0.310 *** (0.083)	-0.313 *** (0.085)
Race/Ethnic Profile (Reference=White)						
Int. Student				0.687 *** (0.126)	0.568 *** (0.135)	0.530 *** (0.140)
Asian				0.112 (0.153)	0.125 (0.151)	0.113 (0.156)
Hispanic				-0.499 *** (0.132)	-0.468 *** (0.132)	-0.494 *** (0.136)
African Americans				-0.646 *** (0.148)	-0.612 *** (0.148)	-0.633 *** (0.152)
Other				-0.531 *** (0.141)	-0.521 *** (0.14)	-0.554 *** (0.144)
Financial Indicators						
Paid with Aid (Any Portion of Tuition)					0.643 *** (0.114)	0.632 *** (0.115)
Paid with Self (Any Portion of Tuition)					-0.145 † (0.085)	-0.151 † (0.086)
Percent Tuition Paid using Aid					-0.611 *** (0.151)	-0.618 *** (0.153)
Pell or Tap recipient					-0.263 ** (0.098)	-0.275 ** (0.101)
Time (Semesters)	0.010 (0.018)	0.049 ** (0.018)	-0.122 *** (0.017)	-0.129 *** (0.017)	-0.132 *** (0.017)	-0.135 *** (0.018)
Constant	4.897 *** (0.095)	5.361 *** (0.102)	-0.172 (0.141)	-0.995 *** (0.231)	-0.723 ** (0.253)	-0.751 ** (0.261)

Note: Coefficients (and standard errors) reported, †p<0.1, *p≤0.05, **p≤0.01, ***p≤0.001

Table C.3.

Students' Cumulative GPA by Peer Association, Additional Behaviors, Academic, Demographic, and Financial Indicators (Models 1 through 5, 5r-limited to students who took at least three courses)

Y=Cumulative GPA	Model 1	Model 2	Model 3	Model 4	Model 5	Model 5r
Behavioral Indicators						
Cumulative Peer Associations (Encountered 2 times)	-0.001 (0.001)	0.000 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Cumulative Peer Associations (Encountered 3 times)	-0.002 (0.002)	0.000 (0.002)	0.002 (0.003)	0.002 (0.003)	0.001 (0.003)	0.001 (0.003)
Cumulative Peer Associations (Encountered 4+ times)	0.003 (0.002)	0.005 * (0.002)	0.007 *** (0.002)	0.006 ** (0.002)	0.006 ** (0.002)	0.006 ** (0.002)
Cumulative Late Registration		-0.082 ** (0.03)	-0.069 * (0.03)	-0.057 † (0.03)	-0.059 * (0.03)	-0.054 † (0.031)
Change Schedule		-0.036 *** (0.011)	-0.032 ** (0.011)	-0.034 ** (0.011)	-0.036 *** (0.011)	-0.035 *** (0.011)
Cumulative Return from Stopout		0.227 *** (0.037)	0.194 *** (0.045)	0.207 *** (0.045)	0.236 *** (0.046)	0.233 *** (0.046)
Academic Indicators						
Ever Need Writing (1st day of semester)			-0.178 *** (0.047)	-0.151 *** (0.045)	-0.150 *** (0.046)	-0.181 *** (0.045)
Ever Need Math (1st day of semester)			-0.400 *** (0.048)	-0.319 *** (0.047)	-0.314 *** (0.047)	-0.313 *** (0.046)
Total Credits to date			0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)
Cumulative Credits attempted			-0.006 *** (0.002)	-0.005 ** (0.002)	-0.006 *** (0.002)	-0.006 *** (0.002)
Cumulative Number of Classes			0.010 † (0.005)	0.010 † (0.005)	0.011 * (0.006)	0.013 * (0.006)
Cumulative Semesters Enrolled as part time			-0.012 (0.015)	-0.012 (0.014)	-0.031 * (0.015)	-0.026 † (0.015)
Demographic Indicators						
Age				0.023 *** (0.003)	0.023 *** (0.003)	0.029 *** (0.003)
Male				-0.252 *** (0.044)	-0.256 *** (0.044)	-0.257 *** (0.043)
Race/Ethnic Profile (Reference=White)						
Int. Student				0.255 *** (0.05)	0.244 *** (0.05)	0.231 *** (0.05)
Asian				0.029 (0.082)	0.027 (0.082)	0.003 (0.081)
Hispanic				-0.343 *** (0.07)	-0.327 *** (0.07)	-0.310 *** (0.074)
African Americans				-0.499 *** (0.078)	-0.483 *** (0.078)	-0.504 *** (0.078)
Other				-0.325 *** (0.074)	-0.318 *** (0.074)	-0.310 *** (0.074)
Financial Indicators						
Cum. Semesters Paid with Aid (Any Portion of Tuition)					0.001 (0.007)	0.002 (0.007)
Cum. Semesters Paid with Self (Any Portion of Tuition)					0.049 *** (0.012)	0.045 *** (0.013)
Cumulative Percent Tuition Paid using Aid					-0.018 (0.041)	-0.038 (0.041)
Cumulative Pell or Tap recipient					-0.005 (0.011)	-0.006 (0.011)
Time (Semesters)	-0.011 * (0.006)	-0.013 † (0.008)	0.006 (0.018)	-0.010 ** (0.018)	-0.029 (0.019)	-0.031 (0.019)
Constant	2.341 *** (0.023)	2.360 *** (0.024)	2.561 *** (0.035)	2.305 *** (0.098)	2.315 *** (0.100)	2.263 *** (0.100)

Note: Coefficients (and standard errors) reported, †$p<0.1$, *$p\leq0.05$, **$p\leq0.01$, ***$p\leq0.001$

Table C.4.

Students' Cumulative Credits Earned by Peer Association, Additional Behaviors, Academic, Demographic, and Financial Indicators (Models 1 through 5, 5r-limited to students who took at least three courses)

Y=Cumulative Credits Earned	Model 1	Model 2	Model 3	Model 4	Model 5	Model 5r
Behavioral Indicators						
Cumulative Peer Associations (Encountered 2 times)	0.395 *** (0.013)	0.275 *** (0.012)	-0.007 (0.007)	-0.008 (0.007)	-0.008 (0.007)	-0.009 (0.007)
Cumulative Peer Associations (Encountered 3 times)	0.566 *** (0.023)	0.428 *** (0.022)	0.061 *** (0.012)	0.061 *** (0.012)	0.058 *** (0.012)	0.057 *** (0.012)
Cumulative Peer Associations (Encountered 4+ times)	0.451 *** (0.016)	0.334 *** (0.015)	0.027 ** (0.009)	0.026 ** (0.009)	0.022 * (0.009)	0.021 * (0.009)
Cumulative Late Registration		-3.050 *** (0.214)	-0.957 *** (0.116)	-0.926 *** (0.117)	-0.871 *** (0.116)	-0.876 *** (0.119)
Change Schedule		0.160 (0.100)	-1.717 *** (0.056)	-1.721 *** (0.056)	-1.683 *** (0.056)	-1.692 *** (0.056)
Cumulative Return from Stopout		-13.403 *** (0.283)	-0.322 † (0.191)	-0.321 † (0.191)	-0.562 ** (0.196)	-0.538 ** (0.198)
Academic Indicators						
Ever Need Writing (1st day of semester)			-0.225 † (0.133)	-0.220 (0.135)	-0.187 (0.135)	-0.208 (0.139)
Ever Need Math (1st day of semester)			0.180 (0.132)	0.154 (0.136)	0.155 (0.136)	0.134 (0.140)
Total Credits to date			0.039 *** (0.002)	0.039 *** (0.002)	0.042 *** (0.002)	0.042 *** (0.003)
Cumulative Credits attempted			0.754 *** (0.009)	0.755 *** (0.009)	0.750 *** (0.009)	0.747 *** (0.009)
Cumulative Number of Classes			0.163 *** (0.027)	0.162 *** (0.027)	0.219 *** (0.028)	0.232 *** (0.028)
Cumulative Semesters Enrolled as part time			-0.232 *** (0.069)	-0.235 *** (0.069)	-0.102 (0.072)	-0.081 (0.073)
Demographic Indicators						
Age				0.018 * (0.009)	0.016 † (0.009)	0.018 † (0.009)
Male				-0.372 ** (0.127)	-0.375 ** (0.127)	-0.352 ** (0.131)
Race/Ethnic Profile (Reference=White)						
Int. Student				0.140 (0.167)	0.091 (0.17)	0.096 (0.175)
Asian				-0.071 (0.235)	-0.024 (0.235)	-0.054 (0.242)
Hispanic				0.103 (0.202)	0.046 (0.204)	0.029 (0.210)
African Americans				-0.129 (0.226)	-0.204 (0.228)	-0.214 (0.234)
Other				-0.278 (0.215)	-0.311 (0.215)	-0.344 (0.223)
Financial Indicators						
Cum. Semesters Paid with Aid (Any Portion of Tuition)					0.097 *** (0.029)	0.096 *** (0.029)
Cum. Semesters Paid with Self (Any Portion of Tuition)					-0.012 (0.063)	-0.021 (0.063)
Cumulative Percent Tuition Paid using Aid					0.546 *** (0.150)	0.543 *** (0.155)
Cumulative Pell or Tap recipient					-0.791 *** (0.078)	-0.806 *** (0.079)
Time (Semesters)	5.632 *** (0.084)	6.915 *** (0.089)	-0.164 † (0.088)	-0.175* (0.088)	0.017 (0.092)	0.002 (0.093)
Constant	-2.369 *** (0.147)	-2.911 *** (0.138)	-0.130 (0.104)	-0.355 * (0.29)	-0.718 * (0.302)	-0.705 * (0.314)

Note: Coefficients (and standard errors) reported, †p<0.1, *p≤0.05, **p≤0.01, ***p≤0.001

Table C.5.

Student's Semestrial Odds of Retention or Graduation by Peer Association, Additional Behaviors, Academic, Demographic, and Financial Indicators (Models 1 through 5, 5r-limited to students who took at least three courses)

Y= Odds of Retention (or Graduation)	Model 1	Model 2	Model 3	Model 4	Model 5	Model 5r
Behavioral Indicators						
Peer Associations (Encountered Twice)	1.184 *** (0.012)	1.168 *** (0.011)	1.020 † (0.011)	1.020 † (0.011)	1.020 † (0.011)	1.021 † (0.011)
Peer Associations (Encountered 3 times)	1.074 *** (0.018)	1.069 *** (0.017)	1.004 (0.017)	1.001 (0.016)	1.001 (0.016)	1.000 (0.016)
Peer Associations (Encountered 4+ times)	1.129 *** (0.016)	1.119 *** (0.015)	1.015 (0.015)	1.010 (0.015)	1.014 (0.015)	1.011 (0.015)
Late Registration		0.555 *** (0.044)	0.649 *** (0.053)	0.675 *** (0.055)	0.729 *** (0.060)	0.768 ** (0.065)
Change Schedule		0.787 *** (0.035)	0.692 *** (0.032)	0.711 *** (0.033)	0.707 *** (0.033)	0.695 *** (0.033)
Return from Stopout		0.568 *** (0.057)	0.718 *** (0.073)	0.716 *** (0.073)	0.717 *** (0.073)	0.677 *** (0.070)
Academic Indicators						
Need Writing (1st day of semester)			0.851 ** (0.052)	0.869 * (0.052)	0.878 * (0.053)	0.849 * (0.051)
Need Math (1st day of semester)			0.686 *** (0.041)	0.719 *** (0.042)	0.732 *** (0.043)	0.717 *** (0.042)
Total Credits Earned (1st day of semester)			1.014 *** (0.001)	1.014 *** (0.001)	1.017 *** (0.002)	1.016 *** (0.002)
Credits attempted in the semester			0.984 * (0.008)	0.985 * (0.008)	0.977 ** (0.008)	0.975 ** (0.008)
Number of Classes in the semester			1.457 *** (0.043)	1.474 *** (0.044)	1.459 *** (0.043)	1.440 *** (0.045)
Part Time in the semester			0.893 (0.069)	0.916 (0.071)	0.828 * (0.065)	0.822 * (0.074)
Demographic Indicators						
Age				1.007 * (0.003)	1.006 † (0.003)	1.010 † (0.003)
Male				0.886 * (0.043)	0.887 * (0.042)	0.882 * (0.042)
Race/Ethnic Profile (Reference=White)						
Int. Student				1.451 *** (0.122)	1.345 *** (0.12)	1.286 *** (0.116)
Asian				0.819 * (0.075)	0.812 * (0.074)	0.823 * (0.074)
Hispanic				0.712 *** (0.056)	0.733 *** (0.058)	0.744 *** (0.058)
African Americans				0.762 ** (0.067)	0.792 ** (0.069)	0.792 ** (0.068)
Other				0.697 *** (0.058)	0.701 *** (0.058)	0.711 *** (0.059)
Financial Indicators						
Paid with Aid (Any Portion of Tuition)					2.141 *** (0.204)	2.130 *** (0.203)
Paid with Self (Any Portion of Tuition)					1.208 * (0.091)	1.206 * (0.092)
Percent Tuition Paid using Aid					0.508 *** (0.066)	0.515 *** (0.067)
Pell or Tap recipient					0.925 (0.054)	0.856 ** (0.050)
Time (Semesters)	0.930 *** (0.009)	0.964 *** (0.011)	0.931 *** (0.013)	0.936 *** (0.013)	0.932 *** (0.013)	0.919 *** (0.012)
Constant	2.465 *** (0.116)	2.769 *** (0.139)	1.149 (0.135)	1.158 ** (0.19)	1.095 (0.202)	1.210 (0.226)

Note: Coefficients (and standard errors) reported, †p<0.1, *p≤0.05, **p≤0.01, ***p≤0.001

Table C.6.

Student's Cumulative Odds of Retention or Graduation by Peer Association, Additional Behaviors, Academic, Demographic, and Financial Indicators (Models 1 through 5, 5r-limited to students who took at least three courses)

Y= Odds of Retention (or Graduation)	Model 1	Model 2	Model 3	Model 4	Model 5	Model 5r
Behavioral Indicators						
Cumulative Peer Associations (Encountered 2 times)	1.048 *** (0.005)	1.048 *** (0.005)	1.013 * (0.006)	1.014 * (0.006)	1.013 * (0.006)	1.004 (0.006)
Cumulative Peer Associations (Encountered 3 times)	1.076 *** (0.008)	1.075 *** (0.008)	1.040 *** (0.009)	1.036 *** (0.008)	1.035 *** (0.009)	1.029 *** (0.008)
Cumulative Peer Associations (Encountered 4+ times)	1.053 *** (0.005)	1.052 *** (0.005)	1.022 *** (0.006)	1.020 *** (0.005)	1.020 *** (0.006)	1.013 * (0.006)
Cumulative Late Registration		0.633 *** (0.036)	0.604 *** (0.038)	0.645 *** (0.039)	0.640 *** (0.04)	0.670 *** (0.041)
Cumulative Change Schedule		0.862 *** (0.021)	0.744 *** (0.024)	0.765 *** (0.024)	0.753 *** (0.024)	0.763 *** (0.024)
Cumulative Return from Stopout		0.893 (0.096)	1.745 *** (0.297)	1.563 ** (0.251)	1.912 *** (0.333)	1.669 ** (0.280)
Academic Indicators						
Ever Need Writing (1st day of semester)			1.061 (0.068)	1.079 (0.066)	1.080 (0.068)	1.054 (0.066)
Ever Need Math (1st day of semester)			0.691 *** (0.043)	0.738 *** (0.045)	0.739 *** (0.047)	0.733 *** (0.046)
Total Credits to date			0.991 *** (0.002)	0.990 *** (0.002)	0.991 *** (0.002)	0.992 *** (0.002)
Cumulative Credits attempted			1.011 * (0.004)	1.012 ** (0.004)	1.009 * (0.004)	1.010 * (0.005)
Cumulative Number of Classes			1.077 *** (0.015)	1.075 *** (0.014)	1.095 *** (0.016)	1.088 *** (0.016)
Cumulative Semesters Enrolled as part time			0.909 * (0.035)	0.930 * (0.034)	0.850 *** (0.037)	0.890 ** (0.036)
Demographic Indicators						
Age				1.009 * (0.004)	1.009 * (0.004)	1.017 * (0.004)
Male				0.854 ** (0.046)	0.843 ** (0.048)	0.833 ** (0.046)
Race/Ethnic Profile (Reference=White)						
Int. Student				1.400 *** (0.13)	1.131 (0.113)	1.115 (0.113)
Asian				0.909 (0.096)	0.918 (0.100)	0.912 (0.098)
Hispanic				0.732 *** (0.066)	0.767 ** (0.072)	0.776 ** (0.072)
African Americans				0.798 * (0.079)	0.839 † (0.087)	0.833 † (0.085)
Other				0.783 ** (0.075)	0.799 * (0.079)	0.800 * (0.078)
Financial Indicators						
Cum. Semesters Paid with Aid (Any Portion of Tuition)					1.050 * (0.024)	1.053 * (0.023)
Cum. Semesters Paid with Self (Any Portion of Tuition)					1.099 ** (0.035)	1.064 * (0.034)
Cumulative Percent Tuition Paid using Aid					1.055 (0.101)	0.892 (0.087)
Cumulative Pell or Tap recipient					0.870 *** (0.021)	0.884 *** (0.020)
Time (Semesters)	0.851 *** (0.01)	0.909 *** (0.02)	0.746 *** (0.034)	0.755 *** (0.032)	0.732 *** (0.034)	0.743 *** (0.033)
Constant	2.769 *** (0.128)	2.965 *** (0.168)	3.812 *** (0.298)	3.529 *** (0.524)	3.408 *** (0.547)	3.700 *** (0.597)

Note: Coefficients (and standard errors) reported, †p<0.1, *p≤0.05, **p≤0.01, ***p≤0.001

APPENDIX D. An Outline of the Computational Programs Sequence for Peer Association

Appendix D outlines the computational program sequence and the transactional data sources needed to conduct Peer Association Calculations.

(1) Identify students for the analysis and append them to a target file (stu1file.dbf) (This file provides a list of dedicated student IDs to run calculations on)

(2) Import all study semesters and courses for the study timeline, and store them in a standalone file -- stusemrg.dbf (This file contains the universe of all relevant semesters for the study time line, course registrations and all students in the universe file. A subset of those students are the target students from step (1) that will be selected later. This file is derived directly from the transactional files. Unlike 'frozen' datafiles, e.g. census, transactional data contains class placements, date stamps, and several additional components crucial for the analysis)

(3) Exclusions part 1 delete all message courses (0 equated courses, courses that serve as a marker only and do not carry any real coursework or serve as an indication of student groups, etc.) from stusemrg.dbf (This allows to ensure that one removes listings that are not actual courses, and receive a subset of the actual transactional activities needed for the investigation)

(4) Exclusions part 2 run offcamp check program to remove off campus rotations, internships, individual studies etc from stusemrg.dbf. (remove listings of courses that may be listed as a single class, by they are not such-e.g. a listing dedicated to sending a student to an internship, listing of clinical rotations, etc. When one replicates such an analysis, it is crucial to consult with the tested institution(s) to determine their recording practices and create a dedicate exclusion list of courses other then zero courseload/equated credits/messaging)

(5) Identify sections of courses: change length of field section to X in stusemrg.dbf file (set space for classes identification, in the cases of this analysis it was 4 or 5 digits that recognize a unique class. Different institutions would identify sections differently. The basic structure of the

computation looks for unique sequence, hence it could be either numbers or letters, e.g. 01234 or 012A3, ABCD1, etc.). Then, replace all section with string representing unique class (classnbr) (The class number is unique for a class in a semester, and that allows to identify students who take classes together at a point in time, e.g. Jim and John who study in the same course, same classroom, would also have the same class number 01234). One can adapt the program to maintain the original value or replace it as deemed necessary.

(6) Run cr0001 -- this creates a record for each student for each semester regardless of whether they attended (This record takes all students in the target and creates a row that would be populated with their information. Some students would not always enroll, e.g. stopout, and for that reason first a record is created per student per semester. As needed, the record will be populated)

(7) Run crterms -- this takes a revised version of stusemrg.dbf and pull out each individual semester eg a file can be named 201202, it stands for the year YYYY along with the term TT the file is named after the term (This process creates a view of each semester in its own unique file, one can create as many semesters as necessary which could be applicable for institutions with different calendars. This increases the robustness of a replication)

(8) Run courseid-- this counts and identifies for each student in the study the number of courses they took each semester and the actual course number of each section. uses classnbr put in section field earlier and deletes semesters for which students took no courses (The view now creates a map of the students in the target group from (1) and their classes)

(9) Run time -- identifies semester of student in the study e.g. 1,2,3 (each semester is indexed so a search for semestrial peers and cumulative peers can be conducted later, and to map the enrollment time, both calendar and active times)

(10) Run cum1 -- this creates a blank file for every student in the study which will house all their classmates. This shell is needed to store all calculations that follow, and to create documentation for each student record. (The product of this process is a file per student, which would hold the documentation of their classmates. Each student record will have a documentation of the classes they took as well as their peer associates)

(11) Run cum2 -- this will populate individual students files created in cum1 with all classmates for entire study time.it will also calculate classmates and unique classmates for each semester for students in study (This core run requires special attention: it takes up to 3 hours to run for up. speed can increase on strong machine, speed can reduce as number of iterations increase. it is recommended to turn on timer at testing, and turn off message system on actual run)

(12) Run cmcountr -- calculates peer frequency for each semester (results in documentation in the target file for the student, in the summary row in the master file, and in an update. The program conducts the loop across all students in the semester, and then moves to the next semester)

(13) Run cmcount1 -- calculates cumulative peer frequency for each point in time (results in documentation of the cumulative peer association at each point in time, e.g. semester 3 for student X, cumulative for semester 3, and in semester 5 for the same student, cumulative for semester 5. The documentation is further updated in the summary file and in the student file)

(14) Conduct posthoc series (checks for students self counts, balances the summaries and ensures that errors are resolved. This process ensure data quality and resolves any transactional data discrepancies that may arise)

(15) Run posthoc1 -- remove the individual herself/himself from the count in each cell. A student cannot be a peer of herself/himself. Then, it repopulate peer counters to fix fields. Summarize revised counts in fix counts.

(16) Run posthoc2 -- conduct a count and match to student file for verification. Verify stopovers and associations. Conducts safety checks and balances. (the two processes ensure the data quality by conducting checks and balances of at multiple levels, e.g. class level, semester level, student cumulative record level, etc; if discrepancies are found they are marked for review)