2014-09-04

A Predictive Model of Canadian College Student Retention

Blair, Morgan

UNIVERSITY OF CALGARY

A Predictive Model of Canadian College Student Retention

by

Morgan Janelle Blair

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF EDUCATION

GRADUATE DIVISION OF EDUCATIONAL RESEARCH

CALGARY, ALBERTA

SEPTEMBER, 2014

Abstract

Estimates in the literature for within-year retention at 2-year colleges range from 57% to 83.9%. This indicates that a large proportion of students who attend 2-year colleges may not be retained beyond the first semester of their studies. Attrition potentially represents a major loss to the student, to the institution, and to society. With current accountability and funding realities becoming more openly discussed, Canadian colleges may not be able to afford to ignore their high rates of attrition in the future.

The focus of this research was to estimate the rate of within-year retention among a sample of students attending two comprehensive community colleges in western Canada, and to develop a predictive model that identified potential determinants of retention among these students. Retention was examined among the total sample, among the sample from each college separately, and among the sample enrolled in each credential type. Astin's Input-Environment-Output model was used as the framework for this research. The model purports that institutional outputs such as retention must be evaluated in the context of the original student inputs and ongoing environmental factors. Multivariable logistic regression was used to develop predictive models of college student retention.

The estimated overall retention rate among this sample was 83.6%, although differences were observed by credential type. Among the aggregate sample, two environmental factors - grade point average and credit load - were the strongest predictors of retention once other factors were considered. The predictors of retention differed by credential type.

The results indicate that the greatest gains in retention may be realized by strategies aimed at encouraging full-time enrolment and supporting academic achievement. The results of

the current study suggest that sub-groups may exist for whom retention is predicted by unique factors. It is important that retention be examined on an institution-by-institution basis. Enhancing our understanding of Canadian college student retention, and taking action to improve retention, may contribute to Canada's future prosperity in a knowledge economy.

Acknowledgements

This dissertation would not have been possible without the help and support of Dr. Peggy Patterson and my committee members: Dr. Dianne Gereluk and Dr. Susan Barker. I'd also like to thank my friends and colleagues who provided support, insight, and feedback whenever it was needed. Special thanks go to my husband, parents, and in-laws, who tirelessly re-arranged their schedules without complaint to accommodate me through this process. Lastly, I would like to thank my son, Hudson, for preventing me from becoming all-consumed by this academic pursuit. Hudson, your vivid imagination and inquisitive innocence kept things in perspective for me.

Table of Contents

List of Tables

List of Symbols, Abbreviations and Nomenclature

Symbol	Definition
CI	Confidence interval
ESL	English as a second language
I-E-O	Input-environment-output
GPA	Grade point average
LR	Likelihood ratio
OR	Odds ratio
PSE	Post-secondary education
SD	Standard deviation
SEM	Strategic enrolment management

CHAPTER 1: INTRODUCTION

Genesis and Rationale

Student retention refers to an academic institution's ability to ensure that students enrol and re-enrol (Astin, 1975; Hagedorn, 2005). Discussions of student retention have permeated academic literature for nearly a century. Among American student samples, within-year (i.e. term-to-term) retention has been shown to be lower among those attending 2-year compared to 4-year colleges (Cofer and Somers, 2000). Estimates in the literature for within-year retention at 2-year colleges range from 57% to 83.9% (Cofer and Somers, 2000; Voorhees; 1987; Hippensteel, St. John, & Starkey, 1996; St. John and Starkey, 1994). In comparison, multiple studies utilizing National Student Aid Survey data estimate the within-year retention at 4-year colleges to be substantially higher, ranging from 87.6%-93.8% (Cofer and Somers, 1997; 1998; Cofer, 1998). This indicates that students who attend 2-year colleges may be at greater risk for attrition than those at 4-year institutions, and that a large proportion of students will not be retained beyond the first semester of their studies. Attrition, then, potentially represents a major loss to the student (in unrealized potential, time, tuition, lost-wages, etc.), to the institution (in revenue, reputation, mission fulfillment, etc.), and to society (as a poor return on investment of public funds) (Canadian Council on Learning, Oct. 2005). However, a gap exists in the research literature on retention at community colleges.

Challenges in Examining Retention

Differences in some of the personal and academic characteristics of university and college students necessitate that retention among these populations be examined separately (Fike & Fike, 2008). Generally, colleges tend to attract higher proportions of 'non-traditional' students

than universities; that is, colleges have a higher proportion of students who are older than age 25 (Fike & Fike, 2008; Aslanian, 2001; Dougherty, 2003), are the first-generation in their families attending a post-secondary institution (Association of Canadian Community Colleges, 2008; Fike & Fike, 2008), belong to a racial minority, study part-time (Cohen and Brawer, 1996; Dougherty, 2003; Fike & Fike, 2008), are more likely to be academically underprepared (McCabe, 2000), and are, or come from families who are of low socioeconomic status (Dougherty, 2003; American Association of Community Colleges, 2000).

Another major difference between college and university students has to do with differences between the types of institutions they attend. Academic programs offered at community colleges in Canada are often shorter than programs offered at universities (i.e. 1-2 year certificate and diploma programs, or 2 year university transfer programs vs. 4 year degree programs). Further, the majority of existing literature on retention is based on American samples. Differences between the American and Canadian systems necessitate that retention be examined from a Canadian perspective for Canadian institutions. As Grayson and Grayson (2003) suggest, "we cannot assume that conclusions based on [American] research are equally applicable to Canada" (p.23). For example, most Canadian colleges are public, whereas in the United States a relatively large proportion of colleges are private. In addition, tuition is often lower at Canadian colleges in comparison with American colleges, and there is also less variation in the quality of education among Canadian compared to American institutions (Grayson, 1998).

Paucity of Canadian research.

Currently there is a lack of consistent research on determinants of retention among Canadian college students, and the majority of the available published research has focused on

American student samples. It is possible that there are differences between American and Canadian samples that may predispose students to be retained or not retained in different ways, thereby reducing the generalizability of the results from American studies to Canadian college students. To date, however, there is no research data which can either confirm or refute this possibility.

Significance

With current accountability and funding realities becoming more openly discussed, Canadian colleges may not be able to afford to ignore their high rates of attrition in the future. In the broader context of the post-secondary system, Canada now faces negative population growth in coming years, and there is no consensus on whether or not increased participation, degree inflation, international recruitment, or immigration will offset these demographic realities (Chagnon & Milan, 2011; Steele, 2010). Since it is more expensive to recruit new students than retain those already enrolled (Joseph, Yakhou, & Stone, 2005), it will become increasingly important to examine institutional retention strategies in order to maintain current enrolment numbers and produce sufficient numbers of graduates to supply labour market demand. However, before such strategies can be developed, institutions should accurately identify, in locally appropriate ways, the characteristics of sub-groups of students who are most likely to be retained, and conversely, those students at greatest risk for attrition.

Purpose

The overarching goal of this research was to produce information that Canadian community colleges can use as a basis for improving the success and the retention of the students that they admit, register and convocate. Therefore, the potential utility of the end results of this

study were considered throughout its design, analysis, and interpretation. With this in mind, the focus of this research was to estimate the rate of within-year retention among a sample of students attending two comprehensive community colleges in western Canada, and to develop a predictive model that identified potential determinants of retention among these students, as well as within each of these two institutions and within each credential. The primary research question for this study was:

1) What are the dynamics of within-year retention among a sample of Canadian community college students?

This research question will be answered through the following sub-questions:

a) What is the estimated rate of within-year retention among Canadian community college students attending two comprehensive community colleges?
b) Are the available data appropriate to develop a predictive model of student retention at two Canadian comprehensive community colleges?
c) What student background and academic characteristics predict within-year retention among those attending these two Canadian comprehensive community colleges?
d) Do the predictive models for retention differ between these two institutions?
e) Do the predictive models differ for students enrolled in university transfer programs, upgrading (and other non-credential programs), and parchment programs leading to a credential?

Institutions that are able to retain at-risk students may be better positioned to compete successfully in a changing educational market (Askari, Lee, Chevannes, Marlin, & Blair, 2011; Black, 2010). The results of this research may be used to identify sub-groups of students who are

most likely to be retained, and conversely, to identify sub-groups of students who are at greatest risk for attrition. The results may have implications for the development of intervention strategies aimed at increasing student retention at a Canadian community college.

Methodology

Astin's input-environment-output (I-E-O) model was used as the framework for this research. According to Astin (1991; 1984), institutional outputs (such as retention, graduates, degrees earned) must be evaluated in the context of the original student inputs and ongoing environmental factors. Input and environmental factors may interact to produce different outcomes, such as whether or not a student is retained.

The I-E-O model was developed to help us understand real world, naturally occurring conditions, and therefore promotes the study of several variables simultaneously (Astin & Antonio, 2012). The I-E-O model is robust in that it requires broad consideration of demographic, social, academic, and financial factors when examining retention, and places equal emphasis on all factors as potential determinants of retention.

Methods

This retrospective secondary analysis of data examined the factors that influence retention among a sample of Canadian college students attending two institutions in western Canada. Data for the 2012-2013 academic year was sourced from the institutional research offices at the two institutions, which routinely collected the data via surveys and administrative records. Students enrolled in English as a Second Language and Apprenticeship programs were excluded. After examining associations between variables using Chi Square tests and binary logistic regression, multivariable logistic regression was used to examine which factors had the

greatest relative impact on retention among the sample as a whole, among the samples attending each college, and among the samples enrolled in university transfer, upgrading/non-parchment, and parchment programs.

Limitations

This research has several limitations, which will be listed briefly here and discussed more thoroughly in Chapter 4. The results of this research may not be generalizable to students other than the cohort used in the analysis due to context-specific circumstances. The results should be interpreted with caution in dissimilar situations. The results of this study are also limited because no differentiation/distinction could be made between students who discontinue studies and those who left to study at another institution (system-level retention). Further, the results do not differentiate between those who dropped out permanently from those students who stopped out temporarily. The manner in which the data were analysed (as categorical as opposed to continuous variables) may mute the effects of some variables on retention. It is also possible that some variables associated with retention were not included in this research. Self-selection bias may have also potentially influenced the results. This research may have also overestimated the rate of retention, since students who are enrolled in the fall semester may drop-out before the add/drop deadline.

Delimitations

There are several delimitations to this research. These include: 1) the fact that data were based on a sample of first-year students, 2) at two community colleges in western Canada, 3) who were enrolled in the first year of a program in the fall semester of the 2012-2013 academic year. These delimitations will be discussed more fully in Chapter 4.

Definitions

Retention:	Student enrolment in the fall *and* subsequent winter semesters at a single institution regardless of program. For the purposes of this paper, retention may be used to in place of 'persistence' to refer to findings in the literature.
Predictive model:	a statistical model developed to estimate the probability of an outcome occurring given certain conditions.
Attrition:	Reduction in a student cohort due to transfers, drop-outs, and stop-outs.
Drop-out:	Permanent withdrawal from post-secondary studies; may be viewed from an institutional or system level.
Stop-out:	Temporary withdrawal from post-secondary studies; may be viewed from an institutional or system level.
Full-time student:	A student who is enrolled in nine or more credit courses in a semester.
Determinant:	A condition, element, or factor that influences the occurrence of a particular outcome.
First generation student:	An individual who is of the first generation in his/her family to enrol in post-secondary education.
University:	The definition may vary by geographic location. However, in the Canadian context it generally refers to institutions offering baccalaureate degrees and/or graduate degrees.

College:	The definition may vary by geographic location. However, in the Canadian context it generally refers to publically funded post-secondary institutions offering broad programming in adult upgrading, certificates, diplomas, university transfer options, and occasionally applied degrees. Collaborative baccalaureate degrees may be offered if need is warranted within the region.
Comprehensive Community College:	The definition may vary by geographic location. In Alberta (the province in which this research is being conducted), see definition of "College" above.
Strategic Enrolment Management	Plans and strategies to make use of data on the student body as actionable information within the context of an institution's vision and mandate.

CHAPTER 2: LITERATURE REVIEW

Differentiating Between Retention and Persistence

Although the focus of this research was on retention, some distinction is required between the terms retention and persistence. Retention and persistence – while often used interchangeably – represent distinctly different perspectives of the same phenomenon. Retention represents an academic institutional measure, and refers to an institution's ability to keep students enrolled from one point in time to another (Astin, 1975; Hagedorn, 2005). Persistence, on the other hand, represents a student measure of academic progress, and refers to a student's motivation to achieve his or her own academic goals. While both perspectives are valuable in their own right, the perspective one takes affects how one goes about researching the topic and the factors considered in that research. For example, research on retention would be conducted from an institutional perspective and examine factors within the scope of institutional control such as admissions requirements. Assuming the student perspective, research on persistence would examine factors relevant to the student such as satisfaction.

While clear definitions for both retention and persistence exist, the terminology is not used consistently in the literature. In fact, persistence and retention are often used interchangeably, which leads to some confusion of the terms (Hagedorn, 2005). The intent of the current research was to examine retention from the institutional perspective. However, because the terminology is often used interchangeably, literature referring to both persistence and retention was referenced to inform the design of this study.

Retention in Context

Before examining the factors that influence student retention, it is important to understand the context in which the Canadian post-secondary education (PSE) system is situated. The importance of retention is underscored by the effects of recent demographic and economic shifts that have occurred in Canada. These shifts, which have shaped the Canadian post-secondary system, have necessitated that institutions focus on retention in order to ensure that sufficient numbers of graduates are produced in order to maintain enrolment and meet labour market demands in the knowledge economy.

Demographics

The demographic shift that is occurring in Canada (and other nations) has and will continue to shape the PSE system. The capacity of the Canadian PSE system was designed around increasing numbers of wealthy baby boomers who participated in higher education in the 1960s and 1970s (Steele, 2010). This ultimately led to the "massification of higher education" (Steele, 2010, p. 28). However, Canada (among other nations) is about to begin experiencing negative population growth (Chagnon & Milan, 2011). This shift will drastically reduce the number of traditional aged students (those aged 18-24) available to participate in PSE in any given year. Statistics Canada data show that since 2006, Canada experienced a higher rate of population growth than did other G8 countries; however the rate of growth (both through immigration and birth) is still declining (Chagnon & Milan, 2011). The effects of negative population growth on the PSE system may be mitigated by a number of factors, including increased participation, degree inflation, international recruitment, and immigration. However, there is no consensus on whether or not these factors will fully compensate for the effects of

negative population growth (Steele, 2010; Statistics Canada, 2010). Changing demographics necessitates that institutions place greater emphasis on retention in order to maintain enrolment numbers and to produce sufficient graduates to support the economy.

The educational needs, expectations, and goals of today's students are also changing, and PSE institutions are struggling to respond to this change (Tagg, 2008). Students are more likely than in the past to delay entry into post-secondary education (Hango, 2011), and may be more likely to take longer to complete their studies than students of the past (Dale, 2010). The current generation of students who form the majority of those enrolled in PSE today (the Millennials) (Howe & Strauss, 2000; Nicoletti & Merriman, 2007) can expect that the kinds of careers that they enter post-graduation will be very different from the careers of generations past: they are more likely to be digitally-based, fast-paced, and require greater depth and breadth of knowledge (Tagg, 2008; Ministry of Training, Colleges and Universities, 2005). This generally means that institutions are required to meet the needs of an increasingly diverse and complex student body, and to ensure that the education students receive prepares them adequately for the labour market post-graduation (Canadian Council on Learning, 2009a). Some provincial ministries overtly recognize that the labour market generally demands greater depth and breadth of knowledge and skill than it has previously, and that this can be achieved through active participation among citizenry in high quality PSE (Rae, 2005).

Economics

Changes in student expectations, goals and needs have accompanied economic changes over time (Tagg, 2008). Globalization has necessitated a more utilitarian, market-oriented approach to PSE that focusses on developing and supporting a knowledge-based economy

(Kirby, 2007; Shanahan & Jones, 2007). This notion is reinforced by literature from the Canadian post-secondary sector over the past decade, which has emphasized the need to produce graduates who can work in a knowledge-based economy (Alberta Enterprise and Advanced Education, 2012; Alberta Advanced Education and Technology, 2007; Canadian Council on Learning, 2009b, 2006; Higher Education Quality Council of Ontario, 2013; Kirby, 2007; Rae, 2005).

At the same time, the economic climate has become increasingly competitive – both in society in general and within the PSE sector specifically (Higher Education Quality Council of Ontario, 2013). Alberta Advanced Education and Technology has recognized that "Alberta's future prosperity will depend on our ability to compete within a global context and advanced education providers have a key role to play in enhancing our competitiveness provincially and locally" (2007, p. 1). In terms of competition, the internet is changing the game of higher education (Steele, 2010). Prospective students now have more options than ever for online flexible programming that meets their unique lifestyles. Accessibility of information, along with the development of online learning will increase competition for students between institutions and countries (Flynn & Vredevoogd, 2010). This presents a challenge and an opportunity for post-secondary institutions since competition for students is now global. Competition will therefore continue to influence the PSE system at the provincial, national, and international levels. In terms of retention, institutions will be competing at a greater level to attract *and* retain students.

Fiscal Realities

Greater competition between PSE institutions has led to more frequent use of accountability frameworks within the Canadian provinces over time, and placed greater emphasis on efficient use of public funds (Alberta Enterprise and Advanced Education, 2012; Higher Education Quality Council of Ontario, 2013; Ministry of Advanced Education, Saskatchewan, 2012; Shanahan and Jones, 2007). While many provinces focus on accountability and efficiency, some provinces, such as Alberta, Ontario, and Quebec, have gone to the extent of adopting performance based funding (Alberta Advanced Education, 2005; Alberta Enterprise and Advanced Education, 2013, September; Council of Ontario Universities, 2010; Pakravan, 2006). Performance based funding mechanisms may be implemented by governments as a means of holding institutions accountable for fulfilling the governments' larger-scale desired outcomes (such as increased graduation rates) (Pakravan, 2006). With current accountability frameworks and performance based funding becoming more openly discussed, it is likely that retention will become increasingly important in the future (Bogue & Dandridge Johnson, 2010).

Retention as an Indicator

Although many performance indicators exist and are used throughout the PSE system, retention tends to be one of the most frequently used quality assurance indicators among institutions (Bogue & Dandridge Johnson, 2010; Tagg, 2008). Institutions often examine performance indicators such as retention rates with the purpose of continuous internal improvement, since once accredited they are primarily responsible for monitoring and improving upon the education that they provide (Canadian Council on Learning, 2009a). Retention is an attractive indicator because the definition of retention used by each institution can depend upon

their unique profile and mission. Goals relating to retention may therefore be set and performance evaluated differently by each institution to suit individual needs (Bogue & Dandridge Johnson, 2010). For example, a small community college offering several one-year certificate programs may choose to define retention at the institutional-level from initial enrolment until the second semester (Hagedorn, 2005). The flexible nature of retention (as an indicator) allows institutions to evaluate performance in ways that are appropriate for their respective missions.

Institutions may also favor the use of retention as a performance indicator because student enrolment and re-enrolment is integral to their function. With current accountability frameworks and performance based funding becoming more openly discussed in several provinces such as Alberta, Ontario, and Quebec (Alberta Advanced Education, 2005; Alberta Enterprise and Advanced Education, 2013; Council of Ontario Universities, 2010; Pakravan, 2006; Shanahan & Jones, 2007), it is likely that retention will become increasingly important in the future (Bogue & Dandridge Johnson, 2010). Fiscal constraints and looming negative population growth will force institutions to pay more attention to strategies to increase retention in order to maintain current enrolment numbers, since it is more expensive to recruit new students than retain those already enrolled (Joseph, Yakhou, & Stone, 2005). Maintaining enrolment numbers may be necessary in order for institutions to remain viable in their current structures. Maintaining enrolment numbers may also be necessary to produce enough skilled graduates to support ministerial objectives relating to the labour market and knowledge based economy (Alberta Enterprise and Advanced Education, 2012; Alberta Advanced Education and

Technology, 2007; Canadian Council on Learning, 2009b, 2006; Higher Education Quality Council of Ontario, 2013; Kirby, 2007; Rae, 2005).

Retention is also an important performance indicator to institutions because it is often used as part of strategic enrolment management (SEM) initiatives. External pressures will increase competition for students between institutions, and therefore the importance of SEM will increase in the future (Smith and Gottheil, 2008). Institutions that shift their focus toward greater emphasis on retention strategies, among other strategies, may have a competitive advantage over others. SEM will be discussed more below.

Calculation of a student retention rate

The calculation of a retention rate is methodologically challenging for community colleges because of the diverse programming offered (e.g. certificates, diplomas, etc.) and complexity of the pathways students often take compared to those seen in universities (Diallo, Trottier, & Doray, 2009). For example, a student may enrol in a one-year certificate program only to switch into a similar two-year diploma program after the first semester. Similarly, a student enrolled in a university transfer program may transfer to university after only one year at college rather than two. The definition of retention for any particular college, then, depends on whether the emphasis is placed on program completion or enrolment from semester-to-semester or year-to-year and whether or not retention should be examined at the institutional or at the program level.

It may not be surprising that a consistently-used definition for the calculation of retention is lacking. Hagedorn (2005) identified four definitions of retention, including system-wide, institutional, discipline or program, and course retention (Table 1 describes the various

definitions as described by Hagedorn). Within each of these definitions (except course retention), the time frame examined can also vary from term-to-term, year-to-year, to entry-to-program completion. Year-to-year institutional-level retention is the most commonly used measure among colleges and universities (Hagedorn, 2005), although this time-frame may not be appropriate for community colleges due to the length of some programs. Issues related to program length will be discussed more below.

Table 1

Definitions of Retention

Definition	Calculation	Pros	Cons
System-wide	Proportion of students who are retained in post-secondary from Time A to Time B regardless of institution.	Tracks students who transfer between institutions to provide a more complete picture of the student life-cycle.	Costly in terms of human and financial resources to establish and maintain database. Requires system-wide compliance which may be difficult to obtain.
Institutional	Proportion of students who are retained from Time A to Time B within an institution.	Provides an overall indicator of institutional effectiveness in relation to retention.	Results may be misleading (especially at colleges) if students must transfer to another institution to complete education.
Discipline or Program	Proportion of students who are retained from Time A to Time B within a specific program at a single institution.	Provides an indicator of program effectiveness.	In smaller programs, data may be too granular to be meaningful and subject to high variability across years.
Course	Proportion of students who are retained from Time A to Time B within a specific course at a single institution.	Provides an indicator of course effectiveness.	Requires an arbitrary cut-point in time at which retention can be determined; data may be granular in small classes; may be better assessed using a D,W,F rate, or credit hours attempted vs. credit hours earned.

Note. Definitions provided in Table 1 are based on definitions provided in Hagedorn, 2005.

Issues with sample selection in previous research

Previous research has tended to use very limited samples when developing predictive models of retention. Selectivity of samples used in traditional calculations of retention may produce misleading results (Pittman, 2008). The majority of research limits inclusion criteria to those who are first-time and/or full-time students (Settle, 2011; Fike & Fike, 2008; Nguyen et al., 2010; Bers, 1988; Wells, 2008), and excludes university transfer students (Hagedorn, 2005). This may systematically bias the results by excluding students who might be considered higher risk for attrition. Such a calculation would also produce a picture of retention that is not representative of the true situation and enrolment mix at most community colleges. Action taken on analyses that do not provide a true picture of retention may result in misinformed enrolment management strategies that adversely affect students and institutions (Green, 2013). Therefore, development of a model based on institutional data that reflects the diversity of the true enrolment mix is required (Hagedorn, 2005).

Given the inconsistency in previous retention research regarding the factors included for analysis, the definitions used, and sample selection; it is not surprising that the results are often incongruent. For the context of Canadian colleges, the paucity of research points to the need for further research focussing specifically on this student population. Further research that adequately represents the enrolment mix at community colleges may be used to develop informed strategies to improve retention.

In a report on higher education in Ontario in 2005, the Ministry of Training, Colleges, and Universities recognized retention as an area requiring further research. The report states that

there is a lack of understanding about why students drop out, and more information on the dynamics of retention would help establish better supports for these students. Following this, the Canadian Council on Learning's (2006) first report on PSE in Canada gave particular focus to improving consistency across jurisdictions on indicators such as retention. Research funded through the Canadian provincial and federal governments has begun to examine reasons for attrition in Canadian PSE (Ma & Frempong, 2008; Dooley, Payne, & Robb, 2011), and manuscripts describing retention at Canadian institutions are increasingly common in the academic literature (this will be discussed more below) (Fisher & Engemann, 2009; Grayson, 1998; Jorgensen, Ferraro, Fichten, & Havel, 2009; Kirby & Sharpe, 2001; Madgett & Belanger, 2008; Sarkar, 1993). However, our understanding of retention at Canadian community colleges is still in its infancy, and further research is warranted.

Factors Related to Student Retention

There is abundant literature identifying factors related to student retention. However, the majority of the available published research has focused on American student samples and there is a lack of consistently identified factors affecting retention among Canadian college students. In the broader literature, these factors have been categorized in a number of ways. Diallo et al. (2009) grouped these factors into financial and non-financial categories, while Herzog (2003) categorizes factors as background, high school preparation, college experience, and financial. Liu (2010) states that factors fall into two categories: those that are external, and those that are academic in nature. Astin and Antonio (2012) conceptualized factors as either input characteristics inherent to the student, or environmental characteristics of the student's experience.

Regardless of the manner in which factors are categorized, previous research has indicated that retention is a multi-faceted issue. The determinants of retention may be constantly evolving and depend on the unique characteristics of each particular cohort of students at each institution. Shifting demographics among post-secondary populations suggest that student retention should be re-examined in new ways in order to maintain a current understanding of the issue (Reason, 2009). This may necessitate that an examination of the determinants of retention consider a broad array of factors.

In order to be inclusive of a broad array of evolving factors already identified, they were categorized as 'input' and 'environmental' in the current study (see Appendix A for a list of some previous research identifying specific factors). Input factors relate to incoming student characteristics, and may include age, gender, direct/delayed entry, parental education, ethnicity, geographic origin, marital status, and dependent children among others. Environmental factors relate to the education context, and may include grade point average (GPA), credit load, friends/family attending the same college, family emotional support, living arrangements, and student loans among others.

Within each of these categories, the results of previous research have shown mixed results. The directionality of association between the independent variables and retention have not always been consistent, and in some cases the results show a lack of association, although divergent results may be partly attributable to differences in sample and methodology. While it is important to keep a broad array of factors in mind when considering retention, focussing on literature specific to studies of Canadian college student samples may provide greater clarity on the issues at hand.

Literature on Canadian college samples

Most of the literature published and cited above is based on American university samples. Canadian literature, although relatively sparse, has focussed mainly on analyses of the *Youth in Transition* survey data collected by Statistics Canada (Parkin and Baldwin, 2009; Finnie and Qiu, 2008; Ma and Frempong, 2008). These analyses have examined retention among postsecondary students in general (as opposed to college and university students independently), and examine system-level retention (Parkin and Baldwin, 2009). Few published studies have examined determinants of retention at Canadian community colleges. The following section will review the few articles describing research that was based on Canadian college samples.

Only four published reports were identified that examined factors associated with retention among Canadian college students (see Appendix B for a detailed description). Among these studies, two were conducted at community colleges (Kirby and Sharpe, 2001; Jorgensen, Ferraro, Fichten, and Havel, 2009) and two were conducted at technical institutes (Fisher and Engemann, 2009; Sarkar, 1993), but were included here due to the paucity of available information. All four of these studies examined factors related to attrition, as opposed to retention.

Jorgensen et al. (2009) examined factors predicting drop-out among 40,682 Canadian college students over ten semesters (attrition rates were examined each semester). They found that low high school grades and older age were the best predictors of attrition. In addition, male gender was also associated with attrition, although to a lesser degree that high school grades and age. Disability status was not found to significantly affect the overall attrition rate, although variations in the patterns of attrition were observed over time between students with and without

disabilities. Students with disabilities were less likely than those without disabilities to drop out between the first and third semesters, but were more likely to drop out in later semesters.

Fisher and Engemann (2009) examined the correlations between year-to-year attrition and five variables among 6,447 college students. These variables included gender, date of admission, program choice, academic preparedness (grades in upgrading), and student engagement (scores on an engagement survey). They found that although attrition was statistically associated with all variables examined, it was predicted best by academic unpreparedness and low levels of student engagement. Academic unpreparedness and low levels of student engagement together explained over half of the variance in enrolment status, indicating that these variables have a strong influence on attrition among this sample.

Kirby and Sharpe (2001) examine factors that influence attrition between the first and second semesters among first year students enrolled in engineering technology programs at a technical institute. They found that those who were retained were significantly more likely than those who were not retained to have higher grades in grade 12 math, have taken a more advanced level of grade 12 math, better high school averages, and to be enrolled in full-time studies. The authors suggest that together, these characteristics represent greater academic integration among those who were retained. Supporting this notion, the results of focus group discussion among a subset of students involved in the study suggest that academic difficulties in the first semester were the main reason behind attrition. Those who were retained also tended to have lower scores on a measure of occupational uncertainty, possibly indicating that they had firmer educational goals. No significant differences were observed between those who were retained and not retained based on gender, age, level of education, or parental education.

Sarkar (1993) collected information from first year students after one-year and two-years of study, and recorded the percentage of students who were retained at each time point. At each time point, the characteristics of students who completed their program vs. those who had not completed their program were compared. The characteristics examined included reasons for taking the program, goal commitment, educational ability, academic/social integration, satisfaction/use of services, student characteristics, and labour market conditions. Sarkar found that completers and non-completers differed on all characteristics except academic/social integration. After one-year, non-completers (those who were not retained) were more likely than those who were retained to have used student support services, to be disabled, female, to be of aboriginal ancestry, married, have dependent children, be employed and work more hours, and be older. Those who were not retained also had lower educational goals, had less understanding of their career choices, were less committed to their educational goals, and had lower previous academic achievement. In addition, non-completers were more likely to have been unemployed when they enrolled in post-secondary studies.

Themes in the Canadian literature

Commonalities exist between the four published Canadian studies examining retention among college students. Factors related to retention that were frequently identified in these reports include past academic performance, engagement/integration, and certain demographic factors. While the methods used to measure each factor may have differed between studies, certain themes exist within the results.

A measure of past academic performance was associated with retention in each of the four reports identified (Fisher and Engemann, 2009; Sarkar, 1993; Kirby and Sharpe, 2001;

Jorgensen, et al., 2009). Academic performance was most often defined in terms of high school GPA, but also as performance in a grade 12 math course, and as more general measures of academic preparedness or ability. Academic preparedness was found to be not only statistically significant in its association with retention, but also of practical significance - it explained a substantial proportion of the variance in enrolment status (Fisher and Engemann, 2009). Therefore, a measure of students' past academic performance may be highly important in predicting whether or not he/she will be retained.

Several studies also confirmed that a measure of student engagement, integration (whether social or academic), or goal commitment may be associated with retention. However, comparison of the results is difficult because each report phrased and/or defined the construct of engagement/integration/goal commitment differently. The constructs used included engagement as measured by a validated survey (Fisher and Engemann, 2009), goal commitment (Sarkar, 1993), academic/social integration (Kirby and Sharpe, 2001), and psychosocial factors (Jorgensen et al., 2009). As with academic preparedness, student engagement was found to be of both statistical significance and practical significance with regards to retention (Fisher and Engemann, 2009). Including a measure of engagement, integration, and/or goal commitment may be an important component of future research examining retention among Canadian college students.

A variety of demographic factors were also shown to be associated with retention in Canadian literature on college samples. Although each study examined different demographic factors, age and/or gender were significantly associated with retention in most of the studies included here. Greater age was inversely associated with retention, so that older students tended

to have a lower rate of retention. The results for gender, however, were mixed. Some results suggest that males (Fisher and Engemann, 2009) – particularly males with a lower high school average – were more likely than females to drop-out (Jorgensen et al., 2009). Conversely, female gender has also been associated with a higher rate of attrition (Sarkar, 1993). Despite the mixed results, gender appears to be associated with retention and its inclusion in future research may help elucidate its role.

While the results of previous research on Canadian college samples confirms several of the factors affecting retention identified in the broader literature, the results are not consistent. This may be partly due to differences in sample selection, methodology, and factors examined. Among the Canadian studies noted above, the factors influencing retention are related to both input and environmental factors. However, inclusion of a greater number of environmental variables such as financial aid/scholarship and academic advising may elicit new information about retention among Canadian college students. In addition, exploration of the role of academic factors such as grade point average and credit load may also prove valuable, particularly among subgroups of students who are retained. Further research among this population is warranted to clarify lingering questions about the determinants of Canadian college student retention.

Why is Retention Important?

Students enter post-secondary education with a variety of aspirations, values, and contexts. Graduation is not always the desired outcome for all students, and certainly in some cases dropping out before graduation is desirable depending on the costs and benefits of competing options (Astin & Antonio, 2012). So why is retention important to understand?

Student retention is an important issue for students, institutions, and society (Canadian Council on Learning, Oct. 2005). The ability to predict which students may be most likely to be retained, and conversely most at risk for attrition, is potentially of value for each of these stakeholder groups for different reasons (Astin, 1975, 1993).

As the direct beneficiaries of post-secondary education (Canadian Council on Learning, 2009a), students have a vested interest in their own college completion. An understanding of the factors that influence retention may help students decide which institution is right for them, and give them a realistic picture of the potential risks and rewards of attendance (Astin, 1993, 1975). The benefits to the individual that stem from post-secondary education may include better health, improved quality of life, and greater lifetime earnings.

Society has a vested interest in post-secondary attainment, as greater attainment promotes both the economic-utilitarian and academic-humanist values that are entrenched in our culture (Kirby, 2007). The economic-utilitarian perspective promotes the value of PSE attainment in economic development (ibid.), while the academic-humanist perspective highlights the role of PSE for personal wellbeing and the collective benefits of society (United Nations, 1948). These values are prominent in PSE sector literature over the last three decades from across the country (Alberta Advanced Education, 2005; Council of Minister of Education, Canada, 1999; Ministry of Advanced Education, Saskatchewan, 2012; Ministry of Education and Training, 1996; Ministry of Training, Colleges and Universities, 2005). For example, Alberta Advanced Education defines a quality advanced education system as one that "…meets the needs of learners, society, and the economy" (2005, p. 3), reflecting both economic-utilitarian and

academic-humanist values. Each provincial government has more or less aligned their objectives and related educational policies to these values over time.

Society as a whole benefits indirectly from greater post-secondary attainment among citizenry (Canadian Council on Learning, 2009a). Corresponding to the economic-utilitarian and academic-humanist values, the benefits of higher levels of post-secondary participation and attainment include greater economic prosperity, productivity, and employment stability, among others. Recent economic and demographic changes require greater participation in post-secondary education in order to meet labour market demand, since a greater proportion of jobs in the future will require a post-secondary education (Alberta Enterprise and Advanced Education, 2013). In addition, greater post-secondary participation and attainment among citizenry will be key to the nation's future prosperity in a knowledge based economy (Kirby, 2007; Shanahan & Jones, 2007).

Institutional administrators have a vested interest in retention risk and the factors contributing to retention. Information on retention can be used to inform educational policy and practices and provide realistic estimates of the potential for impact (Astin, 1993, 1975; Astin & Antonio, 2012). Staff within institutions who are positioned to make decisions that may influence retention include administrators, faculty, admissions officers, student advisors, financial aid staff, student housing staff, and counselors, among others. Ensuring a thorough understanding of retention among individuals in these positions may help the institution as a whole develop sound policies and practices to encourage retention.

The Role of Strategic Enrolment Management

SEM is one vehicle institutions can use to help retain students. By employing strategies to increase retention and using data to inform decision making, institutions can take steps toward ensuring that students enroll, re-enroll, and are prepared to successfully enter the labour market upon graduation.

Enrolment management vs. strategic enrolment management

Enrolment management, according to Hossler and Bean (as cited by McClintock and Snider, 2008), can be defined as "an organizational concept and systematic set of activities designed to enable education institutions to exert more influence over their student enrollments". Enrolment management involves the monitoring of trends among and descriptive statistics on the student body, such as pre-entry characteristics, transfer between institutions, retention, graduate outcomes, and student success (McClintock & Snider, 2008). The addition of the term 'strategic' operationalizes 'enrolment management,' and refers to plans and strategies to make use of data on the student body as actionable information within the context of an institution's vision and mandate. Implementation of strategies derived from the plan relate to the everyday activities of staff as they interact with students and each other to promote the institutions goals.

The role of a SEM plan is to develop focused, informed, and integrated strategies to accomplish the goals set out in the strategic plan as they relate to enrolment management (Hossler & Kalsbeek, 2013). SEM incorporates the entire student experience (Skinkle, 2010), should be informed by historical enrolment, progress, and retention data, recognize capacity, and be placed in context of broader external forces such as the government (Black, 2010). Accordingly, the SEM plan should articulate strategies that bridge recruitment and retention

within the context of a holistic student experience. These strategies ideally form a framework that can be used to go beyond recruitment and retention, to create an evidence-based culture where staff can continually collaborate to provide optimal student experiences. By using data to understand the entire student experience, institutions can guide their practices to better align with student needs, from recruitment through to financial aid and support services, curriculum development, and retention (Henderson, 2012).

Institutions that employ data in enrolment management decision making may become increasingly agile and responsive to changing market conditions (McClintock & Snider, 2008). The need to be agile and flexible in today's market is underscored by pressures to manage costs while growing revenue streams (Black, 2010). Black (2010) states that institutions properly employing SEM will capitalize on external opportunities while moderating threats, make use of data in decision making, and align enrolment strategies with the goals of the institution as a whole. Black (2010) posits that the greatest threat to institutions is that they ignore trends that shift slowly. For example, upward trends in enrolment may lull institutions into a false sense of security, which can often cause leaders to overlook the underlying causes driving enrolment trends in the first place. Data can be used to effectively guide SEM planning if the proper questions are asked of the context (Green, 2013).

The role of data in SEM planning

To be effective, strategic enrolment management must rely on objective, empirical research and data as a guide (Hossler & Kalsbeek, 2013). Yale (2010) makes several recommendations for the kind of data that should be collected for SEM planning. To begin the process, institutions should collect and internally distribute enrolment data, pre-enrolment

characteristics (input data), retention rates, engagement and satisfaction levels, and student outcomes data (Yale, 2010). Where possible, this information should be benchmarked for comparability. Once data of this nature have been collected, they should also be 'drilled down' to examine sub-groups of students (e.g. gender, first generation, Aboriginal). Moreover, institutions should collect and review data from multiple sources, including provincial or national survey data. Yale (2010) suggests specifically examining the following data points when developing SEM strategies: D, W, F rate; retention; mid-term grades; effectiveness of student supports; interventions for academically challenged first year students; redirection of students who stop-out/drop-out.

The types of data that are collected and the research that is conducted in relation to SEM will be most useful to leaders if it is related to the needs of the institution (Skinkle, 2010). Institutions that capitalize on data and research tend to have a research plan integrated into SEM planning, tend to align research priorities with strategic priorities and enrolment targets, and consistently promote data-informed decision making. SEM should not only rely on the descriptive statistics of these data sources, but where possible utilize predictive models of student outcomes to inform decision making (Yale, 2010).

SEM and student outcomes

Data may be viewed as a vehicle for connecting student outcomes (e.g. student success) with input measures (e.g. student characteristics) and environmental factors (e.g. student services, curriculum and teaching) (Shen & Cooley, 2008). Shen and Cooley (2008) argue that, in higher education, the ability of leaders to access, interpret, and act on data needs to be assessed and improved. They stress that decisions should be *informed* rather than *driven* by data;

since data can be used as a foundation for decisions, but responsible decisions also require the consideration of contextual knowledge (e.g. do the data fit and make sense based on our understanding of the situation?). They posit that a common understanding of the role of data is essential for decisions to be data-informed.

Criticisms of SEM

SEM goals and practices have been widely criticized (Hossler & Kalsbeek, 2013). Criticism has often focussed on the potential for SEM goals and practices to undermine the values of higher education. Critics claim that SEM promotes a competitive culture, which may negatively influence access and equity. For example, SEM strategies have been criticized for having too great a focus on increasing admissions selectivity and net revenue, while limiting the opportunities available for students thought to be 'high-risk'.

Certainly, these criticisms may be fair. Institutional approaches to enrolment management often default to the development of strategies aimed at increasing retention rates to the detriment of student access (Green, 2013). In order to uphold values such as access and equity, SEM strategies and practices must be informed by a "calculated, empirical perspective and market orientation to the real costs and consequences of espoused values such as increasing diversity and access or the pursuit of selectivity and prestige" (Hossler & Kalsbeek, pg. 6). If not well-informed by data on the holistic student experience, including data on why and when students are retained, strategies to increase retention may include those that reduce access (such as higher admissions standards) (Hossler & Kalsbeek, 2013).

SEM in Canada

Only 4.6% of Canadian college and university Presidents surveyed about current issues in higher education identified enrolment management as an issue (Patterson, Vanbalkom, Jensen, and Cummings, 2009), despite the fact that SEM planning has been shown to be important to institutional success. However, SEM is less popular in Canada than it is in the U.S. (Smith and Gottheil, 2008). Approximately 62% of institutions in Canada have "adopted a SEM organizational structure", meaning that they have either a committee or a division dedicated to SEM. Interestingly, a committee structure is more prevalent in Canada whereas a divisional structure is more common in the U.S., indicating more advanced SEM systems in the U.S. In Canada the importance of SEM could be considered to be evolving. This may be partly because there is less diversity in quality among Canadian institutions compared to the U.S. (Smith and Gottheil, 2008), so using SEM to compete for students is not as great a priority. Research has shown that only about 28% of Canadian institutions had a comprehensive SEM plan (Smith and Gottheil, 2008). SEM plans at these institutions often included recruitment goals (94%), retention goals (67%), and methods for evaluating and monitoring goal achievement (71%). Approximately half (53%) had effective campus-wide support of SEM initiatives, although the vast majority (89%) felt that retention and success were important issues.

Key SEM trends in Canada

The research mentioned above by Smith and Gottheil (2008) refers to a survey of 274 SEM professionals that examined how SEM was being used at institutions across North America. The remainder of this section will examine some key trends commonly identified in this research.

As mentioned previously, there is substantial interest in examining sub-groups of student potentially at higher risk for adverse outcomes. The most commonly cited sub-groups of interest included Asians, Aboriginals, recent immigrants, rural students, and students with disabilities (Smith and Gottheil, 2008). Smith and Gottheil (2008) also examined retention strategies used at Canadian institutions. The most frequently used retention strategies included disability services, financial aid for emergency situations, Aboriginal services, first-year experience seminars, peer mentoring, consolidated student services and registration, and study skills programs.

Institutions that employ SEM effectively will likely be better equipped to deal with potential declines in enrolment. Declining enrolment in traditional programs does not have to mean a potential collapse of the higher education system if SEM is employed properly. Institutions that shift their focus toward greater emphasis on retention strategies, among other strategies, may be ahead of the curve in terms of SEM planning.

Theories Used to Examine Retention

Research on student retention has previously been based in sociological, psychological, organizational, and interactional theory (see Appendix C for a review of sources that used various theories of student retention). This section will discuss in detail several of the theories used to examine student retention. A synthesis of the commonalities and differences between the theories will then be provided.

Bourdieu's social capital theory

According to Bourdieu (1986), social capital refers to aspects of social structure that people may use as resources to further their interests (Coleman, 1988). For students in post-secondary education, this may include family, friends, instructors, and staff among others.

Closely linked to social capital is cultural capital, which includes artefacts and symbolic indicators of wealth that delineate social class and are passed from generation to generation. Social and cultural capital are often difficult to separate, and are highly linked (and exchangeable with) other forms of capital such as human and financial capital. Capital can be converted from one form to another, and in education this is directly related to reproduction of social classes via accessibility (Bourdieu and Passeron, 1977). Research indicates that students with relatively high levels of cultural and social capital may be more likely to feel entitled to receive a post-secondary education (McDonough, 1997). Accordingly, such students may also have greater resources available to them compared to students with low social/cultural capital, and therefore may be better equipped to achieve their academic goals (Berger, 2000). Berger (2000) theorized that "students with higher levels of cultural capital are more likely to persist, across all types of institutions, than are students with less access to cultural capital" (p. 114). In addition, demographic factors such as age and gender may influence one's level of capital and thus should also be considered when examining retention (Wells, 2008).

Tinto's student integration model

Tinto (1975, 1993) views student retention as a voluntary, longitudinal process. Within this process, inherent traits of the student (e.g. ethnicity, parental support) influence individual levels of commitment to both the institution and to the goal of graduation. Entry traits and commitment influence the degree to which students integrate socially and academically into the institutions. In turn, social and academic integration influence ongoing levels of commitment to the institution and individual goals. Ongoing commitment affects whether or not a student is retained.

Bean and Metzner's student attrition model

Bean and Metzner's (1985, 1990) theory of attrition focuses on the association between student behaviour and retention. Specifically, they consider behaviours to be influenced by a student's attitudes and beliefs, which are in turn influenced by experience with an institution, financial support, and social interaction. Retention is influenced indirectly by these types of experiences. However, this theory considers environmental factors and student intention/goal commitment to be factors that predict student retention. Bean and Metzner see retention as being directly influenced by factors such as a student's demographic characteristics, academic variables, and student satisfaction.

Swail et al.'s model of student retention and achievement

Swail et al.'s (2003) theory suggests that a student's decision to drop-out is influenced by the academic systems of an institution and several cognitive factors, including critical thinking ability, aptitude, learning ability, and time management (among others). These factors are mediated by background characteristics. Within this model, students are connected to the academic system via their respective academic departments.

Astin's input-environment-output model

According to Astin (1991; 1984), institutional outputs (such as retention, graduates, degrees earned) must be evaluated in the context of the original student inputs and ongoing environmental factors. "Inputs refer to the characteristics of the student at the time of entry to the institution; environment refers to the various programs, policies, faculty, peers, and educational experiences to which the student is exposed; and outcomes refers to the student's characteristics after exposure to the environment" (Astin, 1993, p. 7). Input and environmental factors may

interact to produce different outcomes, such as whether or not a student is retained. Retention must be evaluated within the context of the degree of fit (or lack thereof) between the student and the institution. The degree of fit may go beyond demographics to include values, abilities, and goals.

Commonalities and differences between theories

The theories described here share several commonalities, but at the same time they are divergent from one another. Common to each of the theories is the notion that the factors influencing retention can be considered as discrete variables with underlying inter-relationships. For example, several of the theories postulate that incoming student characteristics (e.g. demographics) mediate or even confound the relationship between social relationships and retention. If positioned in a diagram, many of the factors would be associated with each other, as well as retention. This reduces the likelihood that any single factor is responsible for retention. The importance of these inter-relationships is emphasized to a greater extent in certain theories (e.g. Tinto's student integration model, Bordieau's social capital theory) than others.

Many of the theories also broadly position the factors associated with retention as individual or environmental. For example, demographics and social capital are characteristics inherent to the student, while tuition and campus characteristics are inherent to the environment. While this may be an oversimplification of retention, it demonstrates that both student and environmental factors are important to retention.

The theories diverge on the role and responsibility attributed to the student in their academic career. Several of the theories (e.g. Tinto's, Swail et al's, and Bean and Metzner's) view the student as being an active participant in his/her education, although the student may

also be influenced by the surrounding environment. Other theories, such as Astin's and Bordieu's, view the student as taking a more passive role whose fate is mediated by the availability of resources and an interaction with the environment. To an even greater extent, Astin's theory examines retention from an institutional perspective (in terms of outputs), whereas all the other theories examine retention from the student perspective (in terms of the process of reaching goals).

Summary

The contextual factors surrounding retention among Canadian college students must be considered in order to develop a comprehensive understanding of this multi-faceted issue. Demographic changes, economics, and fiscal realities are all part of the dynamics of retention. Published literature on retention has identified a number of factors that are associated with retention, however the determinants of retention may be constantly evolving and depend on the unique characteristics of each particular cohort of students at each institution. As demonstrated by the theories on retention, the determinants of retention may all be inter-related. Understanding the complex relationships between the determinants of retention has value for stakeholders including students, society, and institutional administrators. Among these stakeholders, institutional administrators are perhaps in the best position to make changes that will impact retention – namely through SEM processes.

The current research adds to the literature by providing additional information on the determinants of retention among a sample of Canadian college students. This is one of the first research studies to examine the dynamics of retention among a sample of Canadian students

from more than one community college, and among subgroups of students by credential type. This will add to the relatively sparse literature published on Canadian college student samples.

CHAPTER 3: METHODOLOGY AND METHOD

Theoretical Framework

Astin's input-environment-output (I-E-O) model was used as the framework for this research. The majority of research on retention has positioned the associated factors as being discrete categorical variables with underlying inter-relationships. Positioned this way, it is the presence or absence of these variables and the inter-relationships existing between them over time that determine whether or not students are retained. Astin's model allows for the measurement of discrete input variables while accounting for the mediating effects of environmental variables on outputs such as retention (Astin & Antonio, 2012). Variables that are inherent to the student (i.e. are present at the time of enrolment, such as demographics) may influence the ways in which the student reacts and interacts with environmental factors inherent to the institution (e.g. availability of financial aid).

The I-E-O model was developed to help us understand real world, naturally occurring conditions, and lends itself to complex analyses using multivariate statistics (Astin & Antonio, 2012). Because of this, the model promotes the study of several variables simultaneously. The I-E-O model is robust in that it requires broad consideration of demographic, social, academic, and financial factors when examining retention, and places equal emphasis on all factors as potential determinants of retention.

Astin's model has several benefits for use in retention research. Although the I-E-O model is more simplistic in its conceptualization of retention than other models such as Tinto's, it still describes the interaction between the student and the institution in a meaningful way. Further, the I-E-O model allows for student involvement to be captured by proxy indicators

(such as credits enrolled) rather than subjective retrospective assessments that are expensive and time consuming to collect. The I-E-O model also focusses on naturally occurring phenomena rather than controlled artificial settings, allowing for examination of multiple variables at the same time.

The overarching goal of this research was to produce information that Canadian community colleges can use as a basis for improving the success and the retention of the students that they admit, register and convocate. The focus of this research was to better understand the dynamics of retention among Canadian college students, and in doing so estimate the rate of within-year retention, and to develop predictive models that identified potential determinants of retention among these students, among those attending the two institutions supplying data, and among those enrolled in each credential. The analysis used to answer each of the questions was as follows:

1) **Research Question 1.a.** To estimate the rate of within-year retention among Canadian community college students, the percentage of the total sample that was retained from the fall to winter semester was calculated. In addition, the percentage of students retained at each institution was calculated, as well as the percentage of students enrolled in each credential type who were retained. Chi Square tests were used to test for differences between groups.
2) **Research Question 1.b.** To determine whether or not the data were appropriate for the development of predictive models, the distributions of the data, inter-relationships among variables, and the percentage of missing values in each variable were examined. Power analysis was also conducted for each outcome.

3) **Research Question 1.c.** To identify student characteristics that predict retention among students attending two Canadian community colleges, univariable and multivariable logistic regression was used.

4) **Research Question 1.d.** To determine whether or not the predictive models differ between the two colleges, the reduced multivariable models were compared.

5) **Research Question 1.e.** To determine whether or not the predictive models differ by credential type, multivariable logistic regression was used to develop models based on credential type before the reduced multivariable models were compared.

Methods

Data collection

This secondary analysis used institutional data routinely collected by the institutional research offices at two public, board governed, comprehensive community colleges in Alberta, Canada (College A and College B) from the 2012-2013 academic year. Colleges A and B were selected because they were both willing and able to provide consistent data on a number of factors thought to be associated with retention in the literature. The researcher is also employed at one of the institutions that provided the data. Both College A and B serve under 5,000 credit students annually, and offer a range of programming including certificates, diplomas, university transfer, and trades programs. Demographically, over half of all first year credit students at College A and B were female (64.9% vs. 60.5%, respectively). The average age was 25.3 years at College A, and was 27 years at College B.

Table 2 provides the data sources, variables, and timing of data collection. All data was sourced from voluntary institutional surveys and student records, routinely collected and

compiled by the institutional research offices at each of the Colleges participating in this research. Institutional research staff at each institution prepared the data files and stripped them of identifiers prior to providing them to the researcher for analysis. An indicator variable was created to demarcate each institution before the datasets were appended to form a single file. The combined data set was stored on a password protected computer in the researcher's locked office.

Quality assurance checks were conducted on the combined data file to ensure consistency. During this process it was noted that upgrading students attending College B were missing GPA, and instead were assigned a percentage grade. In these cases, GPA was computed using a translation table to convert percentage grades into GPA. The translation table was provided by College B.

Table 2

Institutional Data Used in a Secondary Analysis Predicting Retention

Survey Data		
• Dependent children • Pass grade 12 (pure) math • Friends attending same institution • Goal commitment • Institutional commitment • Obtained student loan	• Family attending same institution • Expected hours studying • Expected hours working • Academic aspirations • Living arrangements • Passed grade 12 math	• Family emotional support of college attendance • Family financial support of college attendance • First generation student • High school average
Student Records **(***Retrieved after the Add/Drop Deadline)*		
• Date of birth/Age • Gender • Marital Status	• Fall GPA • Aboriginal status • Received loans	• Credit load (full vs. part time) • Credential

Sample

This research used data from a sample of students enrolled in the first year of their academic programs (with exceptions as described below) at Colleges A and B. Students' data was eligible for inclusion in this research if the students were enrolled in the first year of a program and attended new student orientation in the fall semester of 2012, the point at which the majority of data was collected. Therefore census sampling was employed.

Consistent with previously used exclusion criteria (Bers, 1988), students enrolled in a program that takes one semester or less to complete were excluded from the study. This criterion was used because the inclusion of data on students who only intended on completing a one semester program would fall outside the definition of the dependent variable (see below), and their inclusion would therefore produce misleading results. Students who attended new student orientation but who dropped-out before the September add/drop deadline were also excluded, since complete data (including key factors such as fall GPA) were not available for them. Students enrolled in Apprenticeship programs and international students enrolled in ESL were also excluded. Apprenticeship students were excluded because they enrol and attend classes at different times throughout the academic year compared to students in traditional credit programs. International ESL students were excluded because they tend to enrol at the researcher's institution (from which some of the data was drawn) for relatively short-term study-abroad experiences.

Ethical considerations

Ethical approval for the proposed research was sought from the University of Calgary to conduct this research, and also from the Colleges providing data. According to the Tri-council Policy Statement (Article 2.4), this research was exempt from regular ethics review because it relied exclusively on the analysis of secondary data. Therefore, although the research met ethical standards, formal approval was not required. In addition, data sharing agreements were signed between the researcher and both of the Colleges. The data sharing agreements stipulated the information shared and the timeframe in which data sharing occurred. The author is employed at one of the institutions that provided the data.

Dependent Variable

This research examined institutional-level within-year retention. Retention is defined as enrolment in the fall 2012 *and* winter 2013 semesters regardless of program. A student was considered *not* to have been retained (stop-out, drop-out, fail-out) if they enrolled in the fall 2012 but did not attempt to earn any credits in the winter semester of 2013. Retention was determined by obtaining information on student enrolment in the winter semester of 2013.

Institutional-level retention was chosen because this measure was considered by the researcher to be the most useful to the institution and to the researcher. Results and recommendations stemming from an analysis at this level will also have greater potential of being useful to institutions than are other definitions of retention (again, refer to Table 2 for an overview). Within-year (term-to-term) retention was chosen as the time-frame for measurement purposes for several inter-related reasons. First, certificate programs are typically one-year in length when taken as full-time study. Therefore, examining year-to-year retention at the

institutional level (which would include students enrolled in certificate programs who may finish within one year) would produce unrealistically low estimates of retention since the program attendance patterns at the colleges that are part of this study are not developed or intended to last more than one year. Second, at most comprehensive community colleges, certificate programs comprise a substantial proportion of all credit programming. Designing a retention analysis around the program mix available at most community colleges will potentially have greater implications for the utility of the results. Third, descriptive data on historic rates of retention at one of the Colleges in this study shows that a slightly higher proportion of first-year students leave the institution between the fall and winter semesters than between the first and second year of an academic program. Therefore, the first semester may be the critical period for attrition among students at the College. Increasing the understanding of the factors involved in predicting student retention would, therefore, be very useful for both the institutions themselves and for current and future students.

To answer the research questions, the following sub-sets of data were created from the merged data file:

- Aggregate data from students attending both College A and B combined
- Data from students attending College A
- Data from students attending College B
- Data from university transfer students (all university transfer programs)
- Data from students enrolled in parchment programs (certificates, degrees, applied degrees)

- Data from students enrolled in upgrading or other non-parchment programs (academic foundations/upgrading, college preparation, transition programs, general studies, open studies).

Independent Variables

The independent variables used in this research were comprised of a combination of available survey data and variables from student records (see Table 2). Variables with more than two categories and continuous variables were collapsed or dichotomized where appropriate based on the distribution of the data and numbers in each category. Categories within each variable were coded in terms of a reference group believed to be least likely to be retained. As will be described more fully below, this allowed for a comparison of the odds of being retained among each group compared to the group believed to be least likely to be retained.

The independent variables used in this research can also be categorized according to Astin's I-E-O model, as they represent student inputs and environmental factors. Accordingly, student input characteristics were measured by dependent children, grade 12 math, academic aspirations, goal commitment (importance of graduation), institutional commitment (made the right choice in attending college), first-generation student status, age, gender, marital status, aboriginal status, and high school average. Environmental characteristics were measured by friends attending the same institution, family attending the same institution, hours spent studying, hours spent working, housing, family emotional support, family financial support, student loans, credit load, GPA, credential, and institution (College A or B). The institution of attendance (College A or B) was considered as an independent variable for each of the models that used data from both institutions.

Data Mining

With the advent of increased collection and availability of data, data mining has become a popular means of knowledge discovery (Larose, 2005). Hirji (2001) defines data mining as "the analysis and non-trivial extraction of data from databases for the purpose of discovering new and valuable information, in the form of patterns and rules, from relationships between data elements" (p. 87). Hand, Mannila, and Smyth (2001) refined the definition as "the analysis of (often large) observational data sets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner" (p. 1).

One of the main differences between inferential statistics and exploratory data mining is that statistical analysis typically uses hypotheses to infer predictable patterns from data, whereas data mining focuses on the discovery of knowledge without the bounds of *a priori* hypotheses about a problem (Hirji, 2001). Hand et al. (2001) also note that that data mining is often conducted as secondary analyses of large observational data sets. Secondary analyses, such as those conducted in data mining, can also be different from the original purpose of data collection. On the other hand, statistical analyses are often conducted to confirm or refute specific hypotheses related to research questions. Logistic regression, a predictive modeling technique, is one method that can be utilized for both data mining and statistical analysis, and will be discussed in greater detail later in this Chapter.

Logistic regression has increasingly become a common tool used in research on student retention, whether for data mining or for statistical analysis (e.g. Settle, 2011). Appendix D provides the details of previous research identified by the researcher that has used logistic regression to predict retention. The majority of this research has been based on studies done in

the United States. Also, most of the studies noted were retrospective in nature with large samples (greater than 1000); however the samples were frequently restricted to certain sub-groups of students (e.g. students in a certain discipline). While the majority of the studies used institutional-level year-to-year retention as the outcome of interest, term-to-term retention was also examined. The results of these studies demonstrated that financial, social, academic, and demographics factors were all associated with retention to varying degrees.

Analysis and Interpretation

The analysis and interpretation of the data explored within-year retention, with determinants of retention considered as discrete categories. The research used factors associated with retention in previous research in combination with data-mining techniques to examine the dynamics of retention among a sample of Canadian college students. Consistent with the methods used by Blair (2009), this analysis was comprised of univariable and multivariable stages. The analyses were performed using STATA version 11 for Windows.

Univariable analysis.

Chi-square tests (Petrie and Sabin, 2005) were performed to examine where significant differences exist in the categorical data by group. The Chi-square test determines whether or not two variables are associated within a contingency table. In Chi-square tests, the expected frequency of each cell is the product of the relevant row and column totals divided by the overall total. The assumptions of the Chi-square test state that the expected frequency should not be less than 5. The test statistic then calculates the discrepancy between the expected and observed frequencies in every cell, and if the difference is large, it can be concluded that the two variables

are associated (i.e. the null hypothesis is rejected). The equation for a Chi-square test is as follows:

$X_2 = \Sigma (O-E)^2/E$

In this equation O and E are the observed and expected frequencies in each cell of the table. The test statistic follows the Chi-square distribution and has degrees of freedom (r-1)x(c-1).

For reasons mentioned previously, the results of most previous research may or may not be generalizable to Canadian college samples. Moreover, most research does not adequately describe the relationships between independent variables themselves in addition to relationships between the independent and dependent variables. Therefore, selection of variables for inclusion in binary and multivariable analyses did not rely solely on previous research findings. Prior to analysis, independent variables related to retention and attrition among Canadian college students were examined using a table format. The distribution of each variable selected for analysis (as shown in Table 2) was examined by comparing the retained/attrit groups, and within the sample as a whole. Associations between all covariates and retention were initially assessed, and those that exhibit a significant association or trend ($p \leq 0.25$) in 2x2 (or 2x2x2) tables were selected for further examination (Hosmer & Lemeshow, 2000).

Binary logistic regression.

Those variables exhibiting a significant association or trend with retention in 2x2 tables were entered into binary logistic regression (Petrie and Sabin, 2005) to examine the two-way associations between the covariates and retention. Binary logistic regression was then utilized to examine potential correlates of retention. Covariates showing a significant association ($p \leq 0.05$) were selected for inclusion in a multivariable regression model. In addition, demographic

variables (age and gender) shown to be associated with retention in published literature on Canadian college students were also included regardless of the level of significance. This approach has been used in other similar research (for example, see Fike and Fike, 2008).

Logistic regression analysis provides the $\log_e$ odds of a dichotomous outcome adjusted for one or more covariates. Logistic regression is considered appropriate in analyses where the dependent variable is dichotomous (Cabrera, 1994; Menard, 1995). The strength of logistic regression is that a model may contain multiple variables simultaneously that are measured using different scales (Hosmer and Lemeshow, 2000). The equation used was

$$\text{Log}_e[\text{odds}(Y_x=1)]=\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots\ldots,$$

Where β_0 is constant and the regression coefficients ($\beta_{1,2,3\ldots}$) are influenced by the effect of the related covariate ($X_{1,2,3\ldots}$), after controlling for the other covariates included in the model. In logistic regression, the magnitude of the effect of each coefficient is denoted by the deviation of the regression coefficient from zero, where a coefficient '0' indicates no effect.

Multivariable analysis

Multiple logistic regression analysis using backwards elimination and stepwise selection was used to examine potential determinants of retention among the students from each College, among the sample as a whole, and among each credential type (university transfer, parchment programs, and upgrading/other non-parchment programs). The main difference between binary and multivariable logistic regression is that multivariable logistic regression includes more than one covariate in the model. By doing so, each point estimate accounts for the relative effect of other variables included in the model. Variables that were statistically significant ($p=<0.05$) were

kept in the final (reduced) model, in addition to demographic factors (age and gender) shown in published literature to be associated with retention among Canadian college students. Ninety-five percent confidence intervals were presented to illustrate the relative precision of observed point estimates. This analysis was performed so that the beta coefficients were transformed into odds ratios (ORs) automatically in computer output. This aided in interpretation because of the large number of covariates in the multivariable model.

An OR refers to the odds of being a 'case' in an exposed group divided by the odds of being a 'case' in an unexposed group, and can be calculated from a 2 x 2 table:

	Retention - Yes	Retention - No
Characteristic A- Presence	a	b
Characteristic A - Absence	c	d

From a 2 x 2 tables, the OR can be calculated using the following equation:

$$OR = \frac{a \times d}{b \times c}$$

ORs greater than one indicate that the estimated odds of the outcome of interest occurring are greater for group A compared to B. On the other hand, an OR of less than one indicates that group A has an estimated *decrease* in the odds of the outcome of interest relative to group B.

In multivariable regression, several variable selection procedures exist. Backwards elimination and stepwise selection are considered most useful for logistic regression models that involve many covariates (Hosmer & Lemeshow, 2000). In backwards elimination, all covariates are entered into a model and then those with the largest non-significant p-values are removed sequentially until only statistically significant covariates remain in the model. Stepwise selection begins with no covariates in the model and adds them in order of significance (smallest p-value).

As covariates are added to the model, those that become insignificant as more are added can then be removed. This continues until all potential predictors have been added. Hosmer and Lemeshow (2000) recommend backwards elimination followed by a stepwise forwards selection procedure to allow for examination of a number of models.

Both methods, however, allow for the assessment of confounding or effect modification between variables if there is a change in the remaining coefficients. Confounding occurs when two or more covariates are related to each other as well as to the outcome, and can be assessed by examining models using only the outcome and potential confounders and comparing the crude and adjusted estimates. Effect modification occurs when the effect of one independent variable on the dependent variable varies according to the level of another independent variable and it can be assessed by entering interaction terms into the model. Interaction terms can be derived by multiplying two potentially related covariates.

The Likelihood Ratio Test (LR test) was performed to assess the fit of the final multivariable model. The LR test examines the assumption that the saturated and reduced models are the same (i.e. that the reduced model is not a more adequate fit than the saturated model). The LR test is carried out by subtracting the -2 log likelihood of the reduced model from the saturated model and the resulting statistic follows the Chi Square distribution. Degrees of freedom is the difference between n-k, where n is the number of observations and k is the number of parameters fitted in the reduced model.

The retrospective nature of the data allows causal inferences to be drawn based on observed associations; therefore, the results of this analysis were both exploratory and

inferential. Inferences regarding the likely direction of associations were made based on the research literature and on logical criteria (e.g. Hill's criteria) (in Rothman, 2002).

Sample Size Estimation

For a multiple regression analysis, the general rule to ensure adequate sample size is to have at least 10 subjects for each covariate entered in the saturated (original) model (Hosmer & Lemeshow, 2000). The research included data from 837 individuals, which would allow for up to 84 covariates in the aggregate saturated model (the number of variables appropriate for each institution's models, and for the models based on credential type, were lower).

Power analysis was also used to ensure that the sample was adequate for each analysis. Power is the probability of appropriately rejecting a null hypothesis (Petrie and Sabin, 2005), and therefore should be close to one. Since post-hoc sample size calculations are generally regarded as inappropriate (Petrie and Sabin, 2005; Norman and Streiner, 2000), and since the analyses are exploratory, power calculations for logistic regression were performed to ensure that the sample was adequate for the aggregate model, institution-specific models for each institution, and the three credential-type models. Credit load was used as the covariate of interest for each power calculation.

Limitations

The analysis for this research was based on a single cohort of students beginning their studies at each of two community colleges in western Canada in the same year and during the same academic term. Therefore, the results may not be generalizable to cohorts of students with similar or dissimilar characteristics at other institutions. The results should be interpreted with caution in dissimilar situations. The results of this study are also be limited because no

differentiation/distinction could be made between students who discontinue studies and those who left to study at another institution (system-level retention). Further, the results do not differentiate between those who dropped out permanently from those students who stopped out temporarily. Systematic differences in the predictors of retention may exist between students who drop-out, stop-out, and fail-out. Data on reasons for leaving are lacking on the majority of students who leave their studies prematurely; reducing the sample to those with available data would further limit the analysis.

Categorization of continuous variables (such as age or GPA) may mute the effects of these variables on retention. In other words, the observed strength of the relationship between variables may be stronger (or weaker) if the variable is left in its continuous state. However, given the dichotomous nature of the outcome (i.e. retention or attrition) and the complexity of the models, it was felt that the prudent choice would be to dichotomize continuous variables.

It is possible that some variables associated with retention were not included in this research. The variables selected for inclusion were based on characteristics identified from the results of the previous research on the retention of Canadian and American college and university samples cited earlier; however it is possible that some unknown factors may also influence retention among students at these Canadian community colleges. In other cases, data may not have been available or feasible to collect for certain factors previously shown to be influential (e.g. engagement, placement test scores).

Several forms of bias may have also potentially influenced the results. Some of the data was collected from students who voluntarily completed a survey at new student orientation. Students were offered a chance to win a $50 bookstore gift card as an incentive for completing

the survey. Therefore, students who were less financially stable may be more likely to complete the survey (self-selection bias) (Rothman, 2002). However, the risk of self-selection bias may be mitigated by the fact that students were not aware of the incentive until the end of the survey. Self-selection bias may have also been introduced by systematic difference between those who chose to supply identifying information on the survey and those who did not, or those who attended orientation and those who did not. Historically, orientation is attended by about 95% of first-year students at one of the Colleges, and of those who attend approximately 80% complete the survey and provide identifying information. Past comparative analyses of institutional data have not shown major demographic differences existing between those who attend/do not attend, or those who complete/do not complete the survey.

Due to the self-reported nature of many of the variables, recall bias and social desirability bias may have also influenced the results (Rothman, 2002). Recall bias occurs when information is incorrectly provided due to a lapse in memory. For example in this research students may have over or underestimated high school grades, especially if greater time had elapsed between high school graduation and completion of the survey. Social desirability bias occurs when a study participant responds in a way that is perceived to be more favorable in a public domain. In the current study, social desirability bias may have influenced students to portray themselves as studying more, expressing greater commitment to graduation, etc.

This research may have also overestimated the rate of retention, since students who are enrolled in the fall semester may drop-out before the add/drop deadline. Student who dropped-out prior to the add/drop deadline at each institution were not included in this analysis. For these students, complete data on several key factors (e.g. GPA, credit load) was not available.

However, the number of students who drop-out completely before the add/drop deadline is typically small (<20 per institution), so the effect of their exclusion was likely minimal.

Delimitations

There are several boundaries to this research. These bounds include: 1) the fact that data were based on a sample of first-year students, 2) at two community colleges in western Canada, 3) who were enrolled in the first year of a program in the fall semester of the 2012-2013 academic year.

The sample was delimited to first year students only because students enrolled in and who graduate from one-year certificate programs do not typically continue on into a second year of studies. Therefore, certificate students would necessarily be excluded from the analysis if it were to examine all years of study, making the results less useful. These specific institutions were chosen because of 1) the availability of comprehensive information on a large proportion of first-year students at the institution employing the author (comprehensive information of this nature is not typically collected among large proportions of students at Canadian community colleges), and 2) the willingness of other institutions to provide information to the author. The analysis was delimited to data from a single academic year (as opposed to multiple years) in order to restrain the sample size; too large a sample is likely to produce a higher rate of Type I error (indicating that a result is significant when it is not) when using regression (Petrie and Sabin, 2005), since the power of the test becomes too high. Therefore, to increase the accuracy of results, it is preferable to keep the sample size smaller.

CHAPTER 4: FINDINGS

This chapter presents the findings of the data analyses. After reviewing the purpose of the study, the results are presented beginning with a description of the sample and the data. The results corresponding to each research question are then presented in order. The chapter concludes with a summary of the results. The overarching goal of this research is to produce information that institutions can use as a basis for improving the success and retention of the students that they admit, register and convocate.

Accordingly, retention is examined among six subsets of data:

- Aggregate data from student attending both College A and B combined
- Data from students attending College A
- Data from students attending College B
- Data from university transfer students (all university transfer programs)
- Data from students enrolled in parchment programs (certificates, degrees, applied degrees)
- Data from students enrolled in upgrading or other non-parchment programs (academic foundations/upgrading, college preparation, transition programs, general studies, open studies).

Sample

Data were obtained from a combined total of 837 student files (College A: n=644; College B: n=193). These data were collected from a sample of the incoming cohorts of students

at each institution. The samples represent approximately 44.7% of all incoming students at College A and 15% at College B.

Data Distributions

This section thoroughly describes the distributions of all variables considered in this research. This description is meant to aid the reader's understanding of the data in relation to the results. It also provides an avenue for comparison between the sample used in the current research, and samples used in future research. Since retention should be examined on an institution-by-institution basis, this description may also aid in the reader's evaluation of the generalizability of the results to other cohorts of students. This section is descriptive in nature; the implications of relevant associations between variables on the findings will be discussed further on.

First, a description of the aggregate sample characteristics is provided. After this, differences in characteristics between sub-groups of students are highlighted. This includes a description of students who were retained vs. not retained, those who attended College A vs. College B, and those who were enrolled in each credential type (university transfer, parchment, and upgrading/non-parchment programs).

Descriptive statistics can be found in Tables 3-5. Table 3 stratifies the data by retention status. Table 4 stratifies the data by institution (College A and B), and Table 5 stratifies the data by credential type. Within each table the data are organized by student input and environmental factors, and described in this section in the order in which they appear in the tables.

Aggregate sample characteristics

Overall, the average age of the sample was 23.1 years, and 68.8% were female (see Table 3). Nearly one-third (30.1%) were first-generation students. About 10% of students in this sample self-identified as Aboriginal, although the vast majority of those who self-identified came from College B. The majority of students in the total sample were single (81.5%) and one-fifth had dependent children (19.8%). About 48% of students in the sample aspired to earn at least a bachelor's or applied degree. Over 61% reported that they had passed grade 12 pure math[1], while 64% reported their high school average to be between 61-80%. Nearly all students expressed goal commitment (96.4%) and institutional commitment (94.7%). The average GPA among the total sample was 2.68, and 76.8% were enrolled in full-time studies (defined as enrolment in at least nine credits in the fall semester). Approximately half of the students were enrolled in parchment programs; the remainder were enrolled in nearly equal proportions of university transfer programs and upgrading/other non-parchment programs. About two-thirds of student reported that they would study between six and fifteen hours per week. Over two-thirds (77.7%) stated that they had friends attending the same institution and 18.8% had one or more family members attending the same institution. Nearly all students (92.9%) felt that their families were emotionally supportive of their college attendance. The vast majority (86.4%) felt their families were financially supportive of their college attendance. Approximately equal proportions of students were living with their parent's while attending college (38.2%) or in a rental unit (35.5%). The remainder were either living in college residence or in another setting. About two-

[1] Grade 12 pure math refers to the highest level of math that can be completed in high school.

thirds of students in the sample were not working at a job while attending college, and an equal proportion received a student loan.

Differences by retention status

Several variables differed significantly by retention status (see Table 3). Students who were retained were far more likely to have passed grade 12 pure math (p=0.000), have a high school average between 81-100% (p=0.017), express commitment to the goal of graduation (p=0.039), have a fall GPA above the 50th percentile (p=0.000), enroll in full-time studies (p=0.000), and enroll in a university transfer or parchment program (as opposed to upgrading/other non-parchment programs; p=0.000). There was also a trend for students who were retained to state that their families were emotionally supportive of their college education, however this did not reach statistical significance (p=0.052).

Table 3

Cohort Characteristics: Input and Environmental Factors by Retention Status

	Institution, n (%)			
Input and Environmental Factors	Not Retained	Retained	Total	p-value
Student Input Factors				
Age, mean (SD)	23.5 (8.1)	23.1 (6.9)	23.1 (7.1)	0.9453
Age: proportion above median age	73 (53.3)	349 (49.9)	422 (50.4)	0.463
Gender				
Female	95 (69.3)	481 (68.7)	576 (68.8)	0.884
Male	42 (60.7)	219 (31.3)	261 (31.2)	
First generation student				
Don't Know	14 (10.3)	42 (6.0)	56 (6.7)	0.167
No	80 (58.8)	446 (63.7)	526 (62.9)	
Yes	42 (30.9)	212 (30.3)	254 (30.4)	
Aboriginal (self-reported)	15 (11.0)	69 (9.9)	84 (10.0)	0.697
Marital Status				
Married/Common Law	22 (16.1)	90 (13.0)	112 (13.5)	0.617
Separated/Divorced/Widowed	7 (5.1)	35 (5.0)	42 (5.1)	
Single	108 (78.8)	570 (82.0)	678 (81.5)	
Dependent children	27 (19.7)	138 (19.8)	165 (19.8)	0.980
Academic aspirations				
Don't know	28 (20.4)	119 (17.1)	147 (17.6)	0.231
Upgrading	8 (5.8)	25 (3.6)	33 (4.0)	
Certificate/Diploma	45 (32.9)	209 (30.0)	254 (30.5)	
Degree	42 (30.7)	229 (32.9)	271 (32.5)	
Graduate Degree/Prof. designation	14 (10.2)	115 (16.5)	129 (15.5)	
Passed grade 12 math (pure)	62 (46.6)	440 (64.1)	502 (61.2)	0.000
High school average				
<60%	14 (10.5)	52 (7.6)	66 (8.0)	0.017
61-80%	95 (71.4)	431 (62.6)	526 (64.0)	
81-100%	24 (18.1)	206 (29.9)	230 (28.0)	
Commitment to graduation	127 (93.4)	676 (97.0)	803 (96.4)	0.039
Institutional commitment	127 (92.7)	664 (95.1)	791 (94.7)	0.245
Environmental Factors				
Fall GPA, mean (SD)	1.36 (1.4)	2.94 (0.9)	2.68 (1.1)	0.0001
Fall GPA: above 50th percentile	28 (20.4)	400 (57.1)	428 (51.1)	0.000
Credit load: full-time (>9 credits)	57 (41.6)	582 (83.1)	639 (76.3)	0.000
Credential				
University transfer	27 (19.7)	197 (28.1)	224 (26.8)	0.000
Parchment	59 (43.1)	372 (53.1)	431 (51.5)	
Upgrading & other	51 (37.2)	131 (18.7)	182 (21.7)	
Study hours while in school				
0-5 hours	15 (11.0)	46 (6.7)	61 (7.4)	0.301
6-10 hours per week	41 (30.2)	239 (34.5)	280 (33.8)	
11-15 hours	41 (30.2)	230 (33.2)	271 (32.7)	
16-20 hours	27 (19.9)	132 (19.1)	159 (19.2)	
21+ hours	12 (8.8)	45 (6.5)	57 (6.9)	

Input and Environmental Factors	Institution, n (%)		Total	p-value
	Not Retained	Retained		
Friends attending same institution	101 (74.3)	547 (78.4)	648 (77.7)	0.293
Family attending same institution	26 (18.9)	131 (18.7)	157 (18.8)	0.948
Family is emotionally supportive	122 (89.1)	654 (93.7)	776 (92.9)	0.052
Family is financially supportive	119 (87.5)	598 (86.2)	717 (86.4)	0.679
Living arrangements				
College Residence	29 (21.2)	121 (17.3)	150 (17.9)	0.736
Parent's House	50 (36.5)	269 (38.5)	319 (38.2)	
Rental	46 (33.6)	251 (35.9)	297 (35.5)	
Other	12 (8.8)	58 (8.3)	70 (8.4)	
Working while in school				
Not working	46 (33.8)	261 (37.6)	307 (36.9)	0.207
1-10 hours	33 (24.3)	204 (29.4)	237 (28.5)	
11-19 hours	37 (27.2)	160 (23.0)	197 (23.7)	
20+ hours	20 (14.7)	70 (10.1)	90 (10.8)	
Obtained student loan	47 (35.3)	261 (37.9)	308 (37.5)	0.571

Differences between students attending each institution

Table 4 provides the data stratified by institution. The student samples from Colleges A and B are very different when measured by the variables described here. Most of the variables differed significantly by institution; in fact only one variable did not differ significantly. The only variable that did not differ significantly between institutions was whether or not a student's family was financially supportive of his/her college attendance.

Each student input characteristic differed significantly by institution. Compared to students attending College A, students attending College B were on average older (22.5 vs. 25.2 years, p=0.001), and a higher proportion were female (65.8% vs. 78.8%, p=0.001). Although comparable proportions of students were among the first generation in their families to attend college, a higher proportion of students attending College B did not know their parents level of education (14.1%). Compared to students attending College A, a higher proportion of students attending College B self-identified as Aboriginal (0.3% vs. 42.5%, p=0.000), were married (11.1% vs. 21.6%, p=0.000), and had dependent children (12.9% vs. 42.9%, p-0.000). Students attending College B were also more likely than students attending College A to aspire to earn a certificate or diploma (41.4% vs. 27.2%, p=0.000). However, those attending College B were also less likely than students attending College A to have passed grade 12 pure math (31% vs. 70.1%, p=0.000, were less likely to have a high school average between 81-100% (13.8% vs. 32.2%, p=0.000), reported less commitment to the goal of graduation (93.8% vs. 97.2%, p=0.025), and reported less commitment to the institution (91.7% vs. 95.6%, p=0.032).

Differences between institutions were also apparent among the environmental factors examined. Students attending College A had a higher overall GPA compared to those attending College B (2.82 vs. 2.24, p=0.0001). Those attending College A were also more likely to be

enrolled in full-time studies compared to students attending College B (80.8% vs. 61.7%, p=0.000), to be enrolled in university transfer programs (31.2% vs. 11.4%, p=0.000), and to study more per week. While a greater proportion of students attending College A tended to have friends attending the same institution (79.6% vs 71.2%, p=0.014), those attending College B were more likely to have other family members attending the same institution (30.7% vs 15.2%, p=0.000). Students attending College A were more likely than those attending College B to report that their family is emotionally supportive of their college attendance (94.2% vs. 88.6%, p=0.007). A greater proportion of students attending College A lived in a rental unit compared to students attending College B (37.5% vs. 29.0%, p=0.001). Compared to College A, a higher proportion of students attending College B were not working while attending college (33.2% vs. 49.5%, p=0.001). More students from College A obtained a student loan compared to students attending College B (40.9% vs. 22.5%, p=0.000).

Table 4

Cohort Characteristics: Input and Environmental Factors by Institution

	Institution, n (%)			
Input and Environmental Factors	College A	College B	Total	p-value
Student Input Factors				
Age, mean (SD)	22.5 (6.7)	25.2 (8.1)	23.1 (7.1)	0.0001
Age: proportion above median age	357 (55.4)	65 (33.7)	422 (50.4)	0.000
Gender				
Female	424 (65.8)	152 (78.8)	576 (68.8)	0.001
Male	220 (34.2)	41 (21.2)	261 (31.2)	
First generation student				
Don't Know	29 (4.5)	27 (14.1)	56 (6.7)	0.000
No	419 (65.1)	107 (55.7)	107 (55.7)	
Yes	196 (30.4)	58 (30.2)	58 (30.2)	
Aboriginal (self-reported)	2 (0.3)	82 (42.5)	84 (10.0)	0.000
Marital Status				
Married/Common Law	71 (11.1)	41 (21.6)	112 (13.5)	0.000
Separated/Divorced/Widowed	23 (3.6)	19 (10.0)	42 (5.1)	
Single	548 (85.4)	130 (68.4)	678 (81.5)	
Dependent children	83 (12.9)	82 (42.9)	165 (19.8)	0.000
Academic aspirations				
Don't know	115 (17.9)	32 (16.7)	32 (16.8)	0.000
Upgrading	16 (2.5)	17 (8.9)	33 (4.0)	
Certificate/Diploma	175 (27.2)	79 (41.4)	254 (30.5)	
Degree	225 (35.0)	46 (24.1)	271 (32.5)	
Graduate Degree/Prof. designation	112 (17.4)	17 (8.9)	129 (15.5)	
Passed grade 12 math (pure)	444 (70.1)	58 (31.0)	502 (61.2)	0.000
High school average				
<60%	38 (6.0)	28 (14.9)	66 (8.0)	0.000
61-80%	392 (61.8)	134 (71.3)	526 (64.0)	
81-100%	204 (32.2)	26 (13.8)	230 (28.0)	
Commitment to graduation	623 (97.2)	180 (93.8)	803 (96.4)	0.025
Institutional commitment	614 (95.6)	177 (91.7)	791 (94.7)	0.032
Environmental Factors				
Fall GPA, mean (SD)	2.82 (1.0)	2.24 (1.3)	2.68 (1.1)	0.0001
Fall GPA: above 50th percentile	352 (54.7)	76 (39.4)	428 (51.1)	0.000
Credit load: full-time (>9 credits)	520 (80.8)	119 (61.7)	639 (76.3)	0.000
Credential				
University transfer	202 (31.2)	22 (11.4)	224 (26.8)	0.000
Parchment	335 (52.0)	96 (49.7)	431 (51.5)	
Upgrading & other	107 (16.6)	75 (38.9)	182 (21.7)	
Study hours while in school				
0-5 hours	35 (5.5)	26 (13.7)	61 (7.4)	0.000
6-10 hours per week	199 (31.2)	81 (42.6)	280 (33.8)	
11-15 hours	214 33.5)	57 (30.0)	271 (32.7)	
16-20 hours	141 (22.1)	18 (9.5)	159 (19.2)	
21+ hours	49 (7.7)	8 (4.2)	57 (6.9)	

Input and Environmental Factors	Institution, n (%)		Total	p-value
	College A	College B		
Friends attending same institution	512 (79.6)	136 (71.2)	648 (77.7)	0.014
Family attending same institution	98 (15.2)	59 (30.7)	157 (18.8)	0.000
Family is emotionally supportive	605 (94.2)	171 (88.6)	776 (92.9)	0.007
Family is financially supportive	552 (86.5)	165 (85.9)	717 (86.4)	0.836
Living arrangements				
College Residence	109 (17.0)	41 (21.2)	150 (17.9)	0.001
Parent's House	251 (39.0)	68 (35.2)	319 (38.2)	
Rental	241 (37.5)	56 (29.0)	297 (35.5)	
Other	42 (6.5)	28 (14.5)	70 (8.4)	
Working while in school				
Not working	212 (33.2)	95 (49.5)	307 (36.9)	0.001
1-10 hours	194 (30.4)	43 (22.4)	237 (28.5)	
11-19 hours	158 (24.7)	39 (20.3)	197 (23.7)	
20+ hours	75 (11.7)	15 (7.8)	90 (10.8)	
Obtained student loan	260 (40.9)	48 (22.5)	308 (37.5)	0.000

Differences by credential type

Table 5 presents the data distribution when stratified by credential type. University transfer students, students enrolled in parchment programs, and those enrolled in upgrading/non-parchment programs appear to be distinctly different from each other. In fact, only three variables did not exhibit a significant difference between groups, including whether or not a student's family was financially supportive of his/her college attendance, the number of hours spent working per week, and institutional commitment.

Nearly all student input characteristics differed significantly by group. Compared to students enrolled in university transfer programs, those enrolled in parchment programs and upgrading/non-parchment programs tended to be older (20.4 vs. 24.2 vs. 24.1 years, p=0.0001). Compared to university transfer students, higher proportions of parchment students and upgrading/non-parchment students were first generation students (21.0% vs. 33.6% vs. 34.3%, p=0.000). Higher proportions of parchment students and upgrading/non-parchment students self-identified as Aboriginal compared to university transfer students (9.3% vs. 19.2% vs. 4.0%, p=0.000). Compared to university transfer students, higher proportions of parchment and upgrading/non-parchment students were married (5.4% vs. 15.2% vs. 19.3%, p=0.000), and were more likely to have dependent children (7.2% vs. 20.3% vs. 34.1%, p=0.000). University transfer students were far more likely than parchment students and upgrading/non-parchment students to aspire to earn a graduate degree or professional designation (33.6% vs. 6.5% vs. 6.5%, p=0.000), to have passed grade 12 pure math (92.4% vs. 60.6% vs. 22.5%, p=0.000), and to have a high school average between 81-100% (48.2% vs. 22.9% vs. 14.4%, p=0.000). University transfer students and parchment students were more likely than upgrading/non-parchment students to report commitment to the goal of graduation (97.8% vs. 97.4% vs. 94.7%, p=0.003).

Nearly all environmental characteristics varied by credential type. University transfer students and parchment students tended to have a higher GPA than those enrolled in upgrading/non-parchment programs (2.9 vs. 2.8 vs. 2.2, p=0.0001), and were more likely to be enrolled in full time studies (85.7% vs. 83.1% vs. 43.4%, p=0.000). A greater proportion of students enrolled in university transfer programs attended College A (91.2%) and a greater proportion of students enrolled in upgrading/non-parchment programs attended College B (52.2%). University transfer students were also more likely than parchment students and upgrading/non-parchment students to report studying 21 or more hours per week. University transfer students were more likely than those in the other two groups to have friends attending the same institution (88.8% vs. 72.8% vs. 75.7%, p=0.001), while upgrading students were more likely than university transfer and parchment students to have family attending the same institution (28.2% vs. 17.4% vs. 15.6%, p=0.000). University transfer and parchment students were more likely than upgrading/non-parchment students to report that their families were emotionally supportive of their college attendance (95.5% vs. 94.7% vs. 85.6%, p=0.000). University transfer students were more likely than those in the other two groups to live in college residence (71.4% vs. 20.9% vs. 11.5%, p=0.000). Upgrading students were more likely than those enrolled in university transfer and parchment programs to be unemployed while attending college (41.1% vs. 38.4% vs. 34.4%). Students enrolled in parchment programs were more likely than university transfer and upgrading/non-parchment students to receive a student loan.

Table 5

Cohort Description: Student Input and Environmental Factors by Credential Type

Input & Environmental Factors	Credential Type, n (%) University Transfer Programs	Parchment Programs	Upgrading & Non-Parchment Programs	Total	p-value
Student Input Factors					
Age, mean (SD)	20.4 (4.6)	24.2 (7.7)	24.1 (7.4)	23.2 (7.1)	0.0001
Gender					
Female	136 (60.7)	313 (72.6)	127 (69.8)	576 (68.8)	0.007
Male	88 (39.3)	118 (27.4)	55 (30.2)	261 (31.2)	
First generation student					
Don't Know	7 (3.1)	28 (6.5)	21 (11.6)	56 (6.7)	0.000
No	170 (75.9)	258 (59.9)	98 (54.4)	526 (62.9)	0.000
Yes	47 (21.0)	145 (33.6)	62 (34.3)	254 (30.4)	
Aboriginal (self-reported)	9 (4.0)	40 (9.3)	35 (19.2)	84 (10.0)	
Marital Status					
Married/Common Law	12 (5.4)	65 (15.2)	35 (19.3)	112 (13.5)	0.000
Separated/Divorced/Widowed	7 (3.1)	27 (6.3)	10 (4.9)	42 (5.1)	
Single	204 (91.5)	336 (78.5)	153 (75.0)	678 (81.5)	
Dependent children	16 (7.2)	87 (20.3)	62 (34.1)	165 (19.8)	0.000
Academic aspirations					
Don't know	32 (14.4)	80 (18.6)	35 (19.4)	147 (17.6)	0.000
Upgrading	1 (0.5)	7 (1.6)	25 (13.8)	33 (4.0)	
Certificate/Diploma	10 (4.5)	192 (44.7)	52 (28.7)	254 (30.5)	
Degree	105 (47.0)	123 (28.6)	43 (24.8)	271 (32.5)	
Graduate Degree/Prof. designation	75 (33.6)	28 (6.5)	28 (6.5)	129 (15.5)	
Passed grade 12 math (pure)	206 (92.4)	257 (60.6)	39 (22.5)	502 (61.2)	0.000
High school average					
<60%	8 (3.6)	24 (5.7)	34 (19.5)	66 (8.0)	0.000
61-80%	108 (48.2)	303 (71.5)	115 (66.1)	526 (64.0)	
81-100%	108 (48.2)	97 (22.9)	25 (14.4)	230 (28.0)	
Commitment to graduation	218 (97.8)	419 (97.4)	166 (92.2)	803 (96.4)	0.003
Institutional commitment	211 (94.6)	409 (94.9)	171 (94.7)	791 (94.7)	0.974

Input & Environmental Factors	Credential Type, n (%) University Transfer Programs	Parchment Programs	Upgrading & Non-Parchment Programs	Total	p-value
Environmental Factors					
Fall GPA, mean (SD)	2.9 (0.8)	2.8 (1.1)	2.2 (1.4)	2.68 (1.1)	0.0001
Fall GPA: Above 50th percentile	113 (50.5)	246 (57.1)	69 (37.9)	428 (51.1)	0.000
Credit load: Full-time (>9 credits)	192 (85.7)	358 (83.1)	89 (43.4)	639 (74.3)	0.000
Institution					
College A	202 (91.2)	335 (77.7)	107 (52.2)	644 (76.9)	0.000
College B	22 (9.8)	96 (22.3)	75 (41.2)	193 (23.1)	
Study hours while in school					
0-5 hours per week	7 (3.2)	23 (5.4)	31 (17.2)	61 (7.4)	0.000
6-10 hours per week	64 (29.0)	161 (37.7)	55 (30.6)	280 (33.8)	
11-15 hours per week	81 (36.7)	132 (30.9)	58 (32.2)	271 (32.7)	
16-20 hours per week	45 (20.4)	86 (20.1)	28 (15.6)	159 (19.2)	
21+ hours per week	24 (10.9)	25 (5.9)	8 (4.4)	57 (6.9)	
Friends attending same institution	198 (88.8)	313 (72.8)	137 (75.7)	648 (77.7)	0.000
Family attending same institution	39 (17.4)	67 (15.6)	51 (28.2)	157 (18.8)	0.001
Family is emotionally supportive	213 (95.5)	408 (94.7)	155 (85.6)	776 (92.9)	0.000
Family is financially supportive	197 (88.3)	371 (86.3)	149 (84.2)	717 (86.4)	0.482
Living arrangements					
College Residence	39 (71.4)	90 (20.9)	21 (11.5)	150 (17.9)	0.000
Parent's House	122 (54.5)	126 (29.3)	71 (39.0)	319 (38.2)	
Rental	52 (23.2)	179 (41.6)	66 (36.3)	297 (35.5)	
Other	11 (4.9)	35 (8.1)	24 (13.2)	70 (8.4)	
Working while in school					
Not working	86 (38.4)	147 (34.4)	74 (41.1)	323 (37.8)	0.055
1-10 hours	74 (33.0)	120 (28.1)	43 (23.9)	238 (27.9)	
11-19 hours	51 (22.8)	106 (24.8)	40 (22.2)	200 (23.4)	
20+ hours	13 (5.8)	54 (12.7)	23 (12.8)	93 (10.9)	
Obtained student loan	79 (35.4)	186 (44.1)	43 (24.4)	308 (37.5)	0.000

Research Question 1.a: Retention Rates

To answer the first sub-question, the rate of retention was calculated on each of the six sub-sets of data. The estimated rate of within-year retention among all Canadian community college students sampled in this research was 83.6%. Conversely, the estimated rate of within-year attrition was 16.4%. The retention rate for College A was 83.9%, and for College B was 82.9%. The retention rates for College A and College B were not significantly different ($p=0.755$). When examined by credential type, significant differences between retention rates existed. The estimated rate of retention was highest among student enrolled in university transfer programs (88.0%) and parchment programs (86.3%) compared to upgrading/other non-parchment programs (71.9%) ($p=0.000$). Retention rates are presented in Tables 6 and 7.

Table 6

Within-Year Retention Rates by Institution

	Institution		
Retained	College A	College B	Total
No	104	33	137
	75.9%	24.1%	100%
	16.2%	17.1%	16.4%
Yes	540	160	700
	77.1%	22.9%	100%
	83.9%	82.9%	83.6%
Total	644	193	837
	76.9%	23.1%	100%
	100%	100%	100%

p=0.755

Table 7

Within-Year Retention Rates by Credential Type

	Credential Type			
Retained	University Transfer Programs	Parchment Programs	Upgrading and Non-Parchment Programs	Total
No	27	59	51	137
	19.7%	43.1%	37.2%	100%
	12.1%	13.7%	28.0%	16.4%
Yes	197	372	131	700
	28.1%	53.1%	19.0%	100%
	88.0%	86.3%	72.0%	83.6%
Total	224	431	182	837
	26.1%	51.5%	21.7%	100%
	100%	100%	100%	100%

p=0.000

Research Question 1.b: Suitability of the Data for a Predictive Model

The second sub-question called for an assessment of whether or not the data were appropriate for the development of a predictive model. Assessing whether or not the data were appropriate for the development of predictive models was essential in order to ensure that the analysis could proceed as planned, and that the results would be meaningful. The appropriateness of the data was evaluated using four methods that together provide a comprehensive understanding of the data: examining the distributions of the data, examining inter-relationships among variables, evaluating the percentage of missing values in each variable, and conducting power analyses. The results of each of these methods are described below.

Data Distributions

Prior to analysis, it is important to develop a sense of the frequency distributions of the data (Petrie & Sabin, 2005). Examining frequency distributions can assist with identifying trends in the data prior to conducting an analysis. Distributions were examined for the overall sample. As can be seen in the "Total" column of Table 3, the majority of variables have greater than 10% of responses in each category. However, three variables had a high proportion (>90%) of students responding in a similar manner. Nearly 93% of students felt that their families were emotionally supportive of their college attendance. Similarly, nearly 95% were committed to their institution, and over 96% were committed to the goal of graduating. This clustering of the responses in these variables could reduce the reliability of the results if these variables were very strongly associated with each other and with the dependent variable, particularly if they remain significant in the multivariable models. This will be discussed below. Overall, however, the distributions of the variables were thought to be appropriate for the analysis.

Inter-Relationships among Variables

Understanding associations between covariates may aid with interpretation of the multivariable regression models. Several associations between covariates existed that may be important to the interpretation of the models. Appendix E provides the cross-tabulations and the respective Chi Square p-values for some of the potentially important associations between covariates among the aggregate sample. Cross tabulations were performed to examine where collinearity might occur between covariates prior to building the multivariable models, and to suggest which covariates may act as confounding factors in the exploratory models (collinearity refers to extremely strong associations between independent variables – such that the variables should not be included in a multivariable model together). Since correlation between the covariates and institution and credential type were discussed above, they will not be described here. The data displayed in Tables 4 and 5, and in Appendix E, show that there were many inter-relationships between student input and environmental covariates, as was postulated would be the case by Astin & Antonio (2012) of data pertaining to a natural setting.

Several correlations with age were observed among the aggregate sample. A lower proportion of older students (those above the age of 20) had a fall GPA above the 50th percentile compared to younger students (p=0.041). Nearly 47% of older students had a GPA above the 50th percentile (201/422), compared to younger students (227/415). Older students were also less likely to be first-generation students compared to younger students (100/422 vs. 154/414, respectively, p=0.000). Older students were more likely than younger students to have passed grade 12 pure math (301/412 vs. 201/408, respectively, p=0.000), and a greater proportion of older students reported a high school average between 81-100% compared to younger students (138/418 vs. 92/404, respectively, p=0.000).

The students' fall GPA also exhibited numerous correlations with other covariates. A greater proportion of females compared to males had a fall GPA above the 50th percentile (319/576 vs. 109/261, p=0.000). A lower proportion of students who did not know their parents' levels of education had a GPA above the 50th percentile (19/56, 34%) compared to those who were first generation (276/526, 52%) and who were not first generation (133/254, 52%) (p=0.028). A greater proportion of those who had a GPA above the 50th percentile had also passed grade 12 pure math (290/422, 68%) compared to those with a GPA below the 50th percentile (212/398, 53.4%). Not surprisingly, a greater proportion of those with a GPA above the 50th percentile also reported a high school average between 81-100% (155/422, 37%) compared to those who had a GPA below the 50th percentile (75/400, 19%) (p=0.000).

Gender was also associated with two covariates. A greater proportion of females were first-generation students compared to males (189/576 vs. 65/260, p=0.047). Compared to males, a lower proportion of females had passed grade 12 pure math (176/254 vs. 326/566, p=0.001).

First-generation student status was also correlated with several covariates. Students who did not know their parents' levels of education were least likely to have completed grade 12 pure math (24/54, 44%) compared to those who were first-generation (131/249, 53%) and who were not first-generation (347/517, 69%) (p=0.000). Those who didn't know their parents' levels of education also tended to have a lower high school average compared others. Only 15% (8/54) of those who did not know their parents' levels of education had a high school average between 81-100%, compared to 23% of those who were first-generation students (58/247), and 31% of those who were not first-generation students (164/520) (p=0.002).

Having passed grade 12 pure math was associated with having a higher high school average. Compared to those who did not pass grade 12 pure math, those that did pass were more likely to have a high school average between 81-100% (31/312 vs. 195/498, p=0.000).

A high proportion of students responded positively to questions related to both goal commitment and institutional commitment. Over 97% of students who responded positively to goal commitment also responded positively to institutional commitment. Those who agreed that graduation was important were far more likely than those who did not agree to feel that they made the right choice attending college (781/803 vs 8/30; p=0.000).

Associations between variables were also examined in the subsets of data used to develop each predictive model (in addition to the aggregate sample). These subsets included the data for students enrolled at College A, College B, in university transfer programs, in parchment programs, and in upgrading/non-parchment programs. Important associations for each subset of the data (although not described here) are presented in Appendices F-J.

Many associations existed between the variables used in this research. While these associations are important to consider in the interpretation of the final results, there is little indication that collinearity is an issue. Among the associations observed, collinearity would most likely exist between goal commitment and institutional commitment. If entry of both of these variables into a model resulted in collinearity, one of the two would have to be excluded from the model.

The observed associations between independent variables were generally expected. Astin and Antonio (2012) suggest that data collected in any naturally occurring setting will exhibit some degree of association between variables. Therefore the observed association generally

appear appropriate and are not sufficiently strong to detract from the overall suitability of the data for this research.

Missing Values

When a large proportion of data are missing from any particular variable, or from a particular sub-group within the data, the results of an analysis are less likely to be reliable (Petrie & Sabin, 2005). In addition, the number of observations upon which the analysis is conducted may be reduced (Normal and Streiner, 2000). Depending on the proportion and randomness of missing data, different techniques exist to handle the missing observations. Prior to conducting an analysis, it is important to examine patterns in the data in order to minimize any potential bias resulting from missing values.

The percentage of missing values was generally small in this data set. Table 8 presents the percent of the sample missing each variable; variables that are not displayed in this table had complete responses. Because the percent of missing values was generally small (<5%), imputing missing values may have introduced new error into the data or resulted in biased estimates (Norman & Streiner, 2000). Therefore, the prudent choice was to work with the data as is; missing data were not imputed.

Table 8

Percent of Sample Missing Data

	Percent Missing Values		
Variable	Total Sample	College A	College B
Marital status	0.6	0.3	1.4
Family emotional support	0.3	0.3	0.5
Family financial support	0.9	0.9	0.9
Institutional commitment	0.3	0.3	0.5
Goal commitment	0.6	0.5	0.9
Dependent children	0.3	0.2	0.9
Passed grade 12 math	2.1	1.7	3.2
High school average	2.1	1.6	3.7
Living arrangements	0.2	0.2	--
Expected study hours	1.2	0.9	1.9
Expected work hours	0.7	0.8	--
Academic aspirations	0.3	0.2	0.9
Student loan	2.3	1.2	5.6
Family attend	0.1	--	--
Friends attend	0.3	--	--
First-generation student	0.1	--	0.5

Power analysis

Power is the probability of appropriately rejecting a null hypothesis (Petrie and Sabin, 2005). An analysis should only be conducted if there is a good chance of detecting a true effect. Generally, power should be at least 80% in order to detect a true effect. Power is affected by a number of factors, including sample size, variability in the data, the magnitude of the effect of interest, and the significance level of the test used. The likelihood of detecting a true effect is increased if the sample size is large, there is little variability in the data, the effect of the variable of interest is large, and the significance level selected is larger (e.g. p=0.05 vs. 0.01) . Since post-hoc sample size calculations are generally regarded as inappropriate (Petrie and Sabin, 2005; Norman and Streiner, 2000), and since the analyses are exploratory, power calculations for logistic regression were performed to ensure that the sample was adequate for the aggregate

model, institution-specific models for each institution, and the three credential-type models. Credit load was used as the covariate of interest for each power calculation.

The results of the six power analyses are presented in Table 9. All data sets had power above 82.6%, indicating that all had sufficient power to detect a true effect of credit load on retention. For example, with a sample size of 837 in the aggregate data set, there is 100% power at a significance level of 0.05 to detect a change of 6.74 in the odds of retention among those enrolled in full-time studies compared to those enrolled in part-time studies.

Table 9

Results of Power Analyses

Model	Sample Size	Power	Significance Level	Detectable Change in Odds
Aggregate	837	1.00000	0.05	6.74
College A	644	0.99999	0.05	6.00
College B	193	0.99902	0.05	21.56
University Transfer	224	0.94234	0.05	6.75
Parchment Programs	431	0.98714	0.05	10.87
Upgrading & Non-Parchment	182	0.82606	0.05	2.79

Research Question 1.c: Factors Predicting Within-Year Retention

The third research question asked: what student background and academic characteristics predict within-year retention among those attending two Canadian comprehensive community colleges? To answer this question, both univariable and multivariable logistic regression was performed using the aggregate data set (data from College A and B combined).

Univariable Analysis

The frequencies, unadjusted (crude) OR, 95% confidence intervals (95% CI), and p-values for each covariate significantly associated ($p \leq 0.05$) with retention (including age and gender) in univariable analysis are presented in Table 10. Within the table, the 'n(%)' represents the number and proportion of individuals within each category who were retained. Within each

variable, the reference group can be identified by the OR=1.00, since the odds of retention among the reference group is assumed to be equal to one. While all significant associations (in addition to age and gender) are displayed in the table, only those that are most highly significant are reported below.

In the aggregate model, six covariates were significantly associated with retention in univariable analysis and selected for further examination (in addition to age and gender). Three of these were student input factors and three were environmental factors. Although each of the six covariates increased the estimated odds of retention, three environmental covariates in particular were highly significant. Having a GPA above the 50^{th} percentile, enrolment in full-time studies, and enrolment in either a university transfer program or a parchment program all increased the estimated odds of retention (all ORs $\geq$2.45) and were highly significant (p=0.000).

Table 10

Univariable Estimates of Retention in the Aggregate Model

Covariates	n (%)	Crude OR (95% CI)	p-value
Student Input Covariates			
Gender			
Male	219 (83.9)	1.00	
Female	481 (83.5)	0.97 (0.65, 1.44)	0.884
Age			
≤20	351 (84.6)	1.14 (0.79, 1.66)	1.00
>20	349 (82.7)	1.00	
Grade 12 math			
Passed	440 (87.7)	2.04 (1.40, 2.97)	0.000
Did not pass	247 (77.6)	1.00	
High school average			
<60%	52 (78.8)	1.00	
61-80%	431 (81.9)	1.22 (0.65, 2.29)	.534
>81%	206 (89.6)	2.31 (1.12, 4.78)	0.024
Goal commitment (graduation)			
Yes	676 (84.2)	2.28 (1.02, 5.09)	0.044
No	21 (70.0)	1.00	
Environmental Covariates			
Credential			
University transfer	197 (88.0)	2.84 (1.70, 4.76)	0.000
No credential	131 (72.0)	1.00	
Parchment	372 (86.3)	2.45 (1.61, 3.75)	0.000
Credit Load			
Part-time status	118 (59.6)	1.00	
Full-time status	582 (91.1)	6.92 (4.67, 10.26)	0.000
GPA			
Below 50th percentile	300 (73.4)	1.00	
Above 50th percentile	400 (93.5)	5.19 (3.34, 8.07)	0.000

Note. OR=odds ratio; CI=confidence interval; GPA=grade point average.

Multivariable Analysis

Based on the univariable analysis and relevant associations suggested in the literature, all variables listed in Table 10 were included in the multivariable logistic regression model. Regardless of the level of association between age and retention and gender and retention, both of these input factors were entered into the saturated models in order to adjust for the effects of these covariates in the model. Age and gender were kept in the model throughout the elimination process regardless of their levels of significance. The results presented below are adjusted for age and gender, although the results for these two input covariates are not discussed below unless they were significantly associated with retention in the reduced models.

Because these analyses were exploratory and the models complex in terms of the number of potentially relevant covariates, the results were not stratified to examine effect modification and no attempt was made assess confounding by comparing the crude and adjusted estimates. The frequencies, adjusted OR, 95% CIs, and p-values for each covariate associated with retention in multivariable analysis are presented in Table 11 (this represents the reduced models after the elimination process was employed).

Once all covariates showing a significant association with retention in univariable analysis were entered into the multivariable aggregate model, only two environmental variables remained significantly associated with retention (in addition to age and gender). Both enrolment in full-time studies and GPA above the 50^{th} percentile increased the estimated odds of retention (OR= 5.92 and 4.10, respectively) and were both highly significant (p=0.000 and 0.014, respectively). The results of the likelihood ratio test indicate that the reduced model was a significant improvement in the fit of the model compared to the saturated model (Chi Square=38.53, p=0.000).

Table 11

Multivariable Estimate of Retention in the Aggregate Model[a]

Covariates	n (%)	Adjusted OR (95% CI)	p-value
Student Input Covariates			
Gender			
Male	219 (81.1)	1.00	
Female	483 (68.6)	0.72 (0.45, 1.12)	0.148
Age			
≤20	351 (81.6)	1.35 (0.89, 2.05)	0.155
>20	351 (81.6)	1.00	
Environmental Covariates			
Credit Load			
Part-time status	582 (91.1)	1.00	
Full-time status	120 (54.0)	5.92 (3.91, 8.98)	0.000
GPA			
Below 50th percentile	300 (73.4)	1.00	
Above 50th percentile	400 (93.5)	4.10 (2.57, 6.55)	0.000

Note. OR=odds ratio; CI=confidence interval; GPA=grade point average.
[a]n=837

Research Question 1.d: Predictors of Retention by Institution

The fourth sub-question asked: Do the predictive models for retention differ between these two institutions? To answer this question, predictive models of retention had to be developed for each institution, and then compared. For each institution-specific model, both univariable and multivariable logistic regression was performed on the respective data set (data from College A or B only).

Univariable Analysis

The frequencies, unadjusted (crude) OR, 95% confidence intervals (95% CI), and p-values for each covariate significantly associated ($p \leq 0.05$) with retention (including age and gender) in univariable analysis are presented in Tables 12 (College A) and 13 (College B). Within the tables, the 'n(%)' represents the number and proportion of individuals within each category who were retained. Within each variable, the reference group can be identified by the OR=1.00, since the odds of retention among the reference group is assumed to be equal to one. While all significant associations (in addition to age and gender) are displayed in the table, only those that are most highly significant are reported below.

For College A, six variables were significantly associated with retention in the univariable analysis (in addition to age and gender). One of the covariates was a student input factor, while the remaining five were environmental factors. One input variable and three environmental covariates in particular - having passed grade 12 pure math, having a GPA above the 50^{th} percentile, enrolment in full-time studies, enrolment in either a university transfer or parchment program - all increased the estimated odds of retention (all ORs ≥ 2.41) and were highly significant ($p=0.000$).

For College B, four variables were significantly associated with retention in univariable analysis (in addition to age and gender). One of these was a student input factor while the other three were environmental factors. Two environmental factors were highly significant. Being enrolled in full-time studies, and having a GPA above the 50^{th} percentile all increased the estimated odds of retention (all ORs$\geq$26.4) and were highly significant ($p \leq 0.001$).

Table 12

Univariable Estimates of Retention at College A

Covariates	n (%)	Crude OR (95% CI)	p-value
Student Input Covariates			
Gender			
Male	189 (85.9)	1.00	
Female	351 (82.8)	0.79 (0.50, 1.24)	0.307
Age			
≤20	247 (86.1)	1.35 (0.88, 2.07)	0.172
>20	293 (82.1)	1.00	
Grade 12 math			
Passed	389 (87.6)	2.41 (1.56, 3.71)	0.000
Did not pass	141 (74.6)	1.00	
Environmental Covariates			
Credential			
University transfer	177 (87.6)	3.16 (1.76, 5.67)	0.000
No credential	74 (69.2)	1.00	
Parchment	289 (86.3)	2.80 (1.67, 4.69)	0.000
Credit Load			
Part-time status	74 (59.7)	1.00	
Full-time status	466 (89.6)	5.83 (3.69, 9.20)	0.000
GPA			
Below 50th percentile	215 (73.6)	1.00	
Above 50th percentile	325 (92.3)	4.31 (2.69, 6.90)	0.000
Hours spent working			
0 hours per week	186 (83.9)	2.10 (1.06, 4.13)	0.032
1-10 hours per week	167 (86.1)	1.81 (0.92, 3.57)	0.085
11-19 hours per week	125 (79.1)	1.11 (0.57, 2.15)	0.757
20+ hours per week	58 (77.3)	1.00	
Hours spent studying			
0-5 hours per week	25 (71.4)	1.00	
6-10 hours per week	167 (83.9)	2.09 (0.91, 4.76)	0.080
11-15 hours per week	184 (86.0)	2.45 (1.07, 5.62)	0.034
16-20 hours per week	120 (85.1)	2.29 (0.96, 5.44)	0.062
21+ hours per week	39 (79.6)	1.56 (0.57, 4.28)	0.388

Note. OR=odds ratio; CI=confidence interval; GPA=grade point average.

Table 13

Univariable Estimates of Retention at College B

Covariates	n (%)	Crude OR (95% CI)	p-value
Student Input Covariates			
Gender			
Male	30 (73.2)	1.00	
Female	130 (85.5)	2.17 (0.95, 4.95)	0.066
Age			
≤20	104 (81.3)	0.70 (0.30, 1.60)	0.394
>20	56 (86.2)	1.00	
First-generation status			
First-generation	51 (87.9)	3.64 (1.18, 11.21)	0.024
Not first-generation	91 (85.1)	2.84 (1.09, 7.43)	0.033
Unknown	18 (66.7)	1.00	
Environmental Covariates			
Living arrangements			
College residence	28 (68.3)	0.72 (0.24, 2.11)	0.547
Parent's home	62 (91.2)	3.44 (1.04, 11.41)	0.043
Rental unit	49 (87.5)	2.33 (0.73, 7.49)	0.154
Other	21 (75.0)	1.00	
Credit Load			
Part-time status	44 (59.7)	1.00	
Full-time status	116 (97.5)	26.4 (7.66, 90.79)	0.000
GPA			
Below 50th percentile	85 (72.7)	1.00	
Above 50th percentile	75 (98.7)	28.24 (3.76, 211.67)	0.001

Note. OR=odds ratio; CI=confidence interval; GPA=grade point average.

Multivariable analysis

Based on the univariable analysis and relevant associations suggested in the literature, all covariates listed in Table 12 were included in the multivariable logistic regression model for College A. All covariates listed in Table 13 were included in the multivariable logistic regression model for College B. Regardless of the level of association between age and retention and gender and retention, both were entered into the saturated models in order to adjust for the effects of these covariates in the model. Age and gender were kept in the model throughout the elimination process regardless of their levels of significance. The results presented below are adjusted for age and gender, although the results for these two input covariates are not discussed below unless they were significantly associated with retention in the reduced models.

Because these analyses were exploratory and the models complex in terms of the number of potentially relevant covariates, the results were not stratified to examine effect modification and no attempt was made assess confounding by comparing the crude and adjusted estimates. The frequencies, adjusted OR, 95% CIs and p-values for each covariate associated with retention in multivariable analysis are presented in Table 14 (this represent the reduced models after the elimination process was employed).

After adjusting for all other covariates in the model for College A, only three covariates remained significantly associated with retention in the reduced model. Gender, which was not associated with retention in univariable analysis but was included as a demographic factor, became significant in the reduced model. Female gender decreased the estimated odds of retention (OR=0.57, p=0.030). Two environmental covariates – enrolment in full-time studies and having a GPA above the 50^{th} percentile - both increased the estimated odds of retention (OR=5.34 and 3.88, respectively; both p=0.000). The results of the likelihood ratio test indicate

Table 14

Multivariable Estimates of Retention for Institution-Specific Models

	Model					
	College A[a]			College B[b]		
Covariates	n (%)	Adjusted OR (95% CI)	p-value	n (%)	Adjusted OR (95% CI)	p-value
Student Input Covariates						
Gender						
Male	189 (85.9)	1.00		30 (60.0)	1.00	
Female	351 (82.8)	0.57 (0.34, 0.95)	0.030	132 (79.5)	1.45 (0.52, 4.09)	0.475
Age						
≤20	247 (86.1)	1.38 (0.86, 2.22)	0.183	104 (72.7)	0.64 (0.23, 1.81)	0.407
>20	293 (82.1)	1.00		58 (79.5)	1.00	
Environmental Covariates						
Credit Load						
Part-time	74 (59.7)	1.00		46 (47.4)	1.00	
Full-time	466 (89.6)	5.34 (3.30, 8.65)	0.000	116 (97.5)	19.22 (5.42, 68.13)	0.000
GPA						
Below 50th percentile	215 (73.6)	1.00		85 (72.7)	1.00	
Above 50th percentile	325 (92.3)	3.88 (2.35, 6.42)	0.000	75 (98.7)	18.59 (2.33, 148.02)	0.006

Note. OR=odds ratio; CI=confidence interval; GPA=grade point average.

[a]n=644

[b]n=193

that the reduced model was a significant improvement in fit compared to the saturated model (Chi Square=30.26, p=0.000).

The reduced model for retention at College B contained two environmental covariates that remained significantly associated with retention (in addition to age and gender). As with the other models, enrolment in full-time studies and having a GPA above the 50th percentile both increased the estimated odds of retention (OR=19.22, p=0.000; OR=18.59, p=0.006, respectively). The results of the likelihood ratio test indicate that the reduced model was a significant improvement in fit compared to the saturated model (Chi Square=20.12, p=0.000).

Comparison of the models

The reduced multivariable models of retention for Colleges A and B differed in several ways. These included the covariates that remained significant, the magnitude of the effects detected, and the precision of the estimates (spread of the 95% CIs). Although credit load and GPA were important factors in both models, gender was also a significant predictor among students for College A but not for College B. Specifically, female gender decreased the estimated odds of retention for College A, but was not significantly associated with retention for College B.

Although GPA and credit load predicted retention at both institutions, the actual point estimates for credit load and GPA differed. Table 14 shows that the estimated OR for credit load and GPA for College A are relatively small compared to those for College B. For example, the OR for GPA is 3.88 at College A and is 18.59 at College B. Further, the precision of the estimates also differ between institutions; the precision of the estimates for College B are far less than that for College A. For example, the 95% CI for GPA for College A is 2.35-6.42, whereas for College B it is 2.33-148.02.

Research Question 1.e: Predictive Models by Credential Type

Sub-question 1.e asked: do the predictive models differ for students enrolled in university transfer programs, upgrading and other non-credential programs, and parchment programs leading to a credential? To answer this question, predictive models of retention had to be developed for each credential type, and then compared. For each credential-specific model, both univariable and multivariable logistic regression was performed on the respective data set.

Univariable Analysis

The frequencies, unadjusted (crude) OR, 95% confidence intervals (95% CI), and p-values for each covariate significantly associated ($p \leq 0.05$) with retention (including age and gender) in univariable analysis are presented in Table 15-17. Within the tables, the 'n(%)' represents the number and proportion of individuals within each category who were retained. Within each variable, the reference group can be identified by the OR=1.00, since the odds of retention among the reference group is assumed to be equal to one. While all significant associations (in addition to age and gender) are displayed in the table, only those that are most highly significant are reported below.

Considering only data from students enrolled in university transfer programs, three environmental variables were significantly associated with retention in the univariable analysis (exclusive of age and gender). Being enrolled full-time, having a GPA above the 50^{th} percentile, and working less than or equal to ten hours per week were all associated with increased estimated odds of retention (all $OR \geq 4.16$, all $p \leq 0.016$).

When data from students enrolled in parchment programs was considered, three variables were significantly associated with retention (in addition to age and gender) in univariable analysis. While this included one student input factor, the two environmental factor in particular,

including full-time credit load and having a GPA above the 50th percentile, greatly increased the estimated odds of retention (both OR≥10.53) and were highly significant (p=0.000).

Three variables were significantly associated with retention when only data from students enrolled in upgrading and other non-parchment programs was considered (in addition to age and gender) in univariable analysis. Two of these were student input factors, while one was an environmental factor. The environmental factor (full-time credit load) was most highly significant (OR=2.75, p=0.004). However, two input factors (goal commitment and institutional commitment) also significantly increased the estimated odds of retention (both OR≥3.94, both p≤0.016).

Table 15

Univariable Estimates of Retention in University Transfer Programs

Covariates	n (%)	Crude OR (95% CI)	p-value
Student Input Covariates			
Gender			
Male	75 (85.2)	1.00	
Female	122 (89.7)	1.51 (0.67, 3.34)	0.317
Age			
≤20	53 (88.3)	1.05 (0.42, 2.63)	0.914
>20	144 (87.8)	1.00	
Environmental Covariates			
Credit Load			
Part-time	20 (62.5)	1.00	
Full-time	177 (92.2)	7.08 (2.91, 17.22)	0.000
GPA			
Below 50th percentile	90 (81.1)	1.00	
Above 50th percentile	107 (94.7)	4.16 (1.61, 10.76)	0.003
Hours spent working			
0 hours per week	79 (91.9)	7.05 (1.81, 27.44)	0.005
1-10 hours per week	66 (89.2)	5.16 (1.35, 19.6)	0.016
11-19 hours per week	44 (86.3)	3.93 (1.00, 15.50)	0.051
20+ hours per week	8 (61.5)	1.00	

Note. OR=odds ratio; CI=confidence interval; GPA=grade point average.

Table 16

Univariable Estimates of Retention in Parchment Programs

Covariates	n (%)	Crude OR (95% CI)	p-value
Student Input Covariates			
Gender			
Male	104 (88.1)	1.00	
Female	268 (85.6)	0.80 (0.42, 1.52)	0.499
Age			
≤20	212 (89.1)	1.68 (0.97, 2.92)	0.066
>20	160 (82.9)	1.00	
Grade 12 math			
Passed	230 (89.5)	2.02 (1.16, 3.52)	0.013
Did not pass	135 (80.8)	1.00	
Environmental Covariates			
Credit load			
Part-time	40 (54.8)	1.00	
Full-time	332 (92.7)	10.53 (5.73, 19.38)	0.000
GPA			
Below 50th percentile	132 (71.4)	1.00	
Above 50th percentile	240 (97.6)	16.06 (6.73, 38.36)	0.000

Note. OR=odds ratio; CI=confidence interval; GPA=grade point average.

Table 17

Univariable Estimates of Retention in Upgrading/Non-Parchment Programs

Covariates	n (%)	Crude OR (95% CI)	p-value
Student Input Covariates			
Gender			
Male	40 (72.3)	1.00	
Female	91 (71.7)	0.95 (0.47, 1.92)	0.882
Age			
≤20	86 (73.5)	1.23 (0.63, 2.40)	0.539
>20	45 (69.2)	1.00	
Goal commitment (graduation)			
Yes	124 (74.7)	3.94 (1.29, 12.00)	0.016
No	6 (42.9)	1.00	
Institutional Commitment			
Yes	126 (73.4)	4.2 (1.13, 15.60)	0.032
No	4 (40.0)	1.00	
Environmental Covariates			
Credit load			
Part-time	60 (51.7)	1.00	
Full-time	73 (82.0)	2.75 (1.39, 5.46)	0.004

Note. OR=odds ratio; CI=confidence interval; GPA=grade point average.

Multivariable analysis

Based on the univariable analysis and relevant associations suggested in the literature, all covariates listed in Table 15 were included in the multivariable logistic regression model for university transfer programs. Likewise, the covariates listed in Table 16 were included in the model for parchment programs, and the covariates listed in Table 17 were included in the model for upgrading/non-parchment programs. Regardless of the level of association between age and retention and gender and retention, both of these input factors were entered into the saturated models in order to adjust for the effects of these covariates. Age and gender were kept in the models throughout the elimination process regardless of their levels of significance. The results presented in Table 18 are adjusted for age and gender, although the results for these two input covariates are not discussed below unless they were significantly associated with retention in the reduced models.

Once all covariates were entered into the retention model among university transfer students, only three environmental factors remained significant in the reduced model and all increased the odds of retention. Compared to those who were enrolled part-time, those enrolled in full-time studies had an increased estimated odds of retention (OR=5.11, p=0.001). As with many of the other models, having a GPA above the 50^{th} percentile resulted in over a four-fold increase in retention (OR=4.10, p=0.009). Also, those who were not working, and those working ten or fewer hours per week had an increase in the estimated odds of retention compared to those who worked 20 hours or more per week (OR= 5.87, p=0.023; and OR=4.69, p=0.043, respectively). Because all variables remained significantly associated with retention when

Table 18

Multivariable Estimates for Retention in University Transfer, Parchment, and Upgrading/Non-Parchment Program Models

	Model								
	University Transfer Programs[a]			Parchment Programs[b]			Upgrading & Non-Parchment Programs[c]		
Covariates	n (%)	Adjusted OR (95% CI)	p-value	n (%)	Adjusted OR (95% CI)	p-value	n (%)	Adjusted OR (95% CI)	p-value
Student Input Covariates									
Gender									
Male	75 (85.2)	1.00		104 (88.1)	1.00		40 (72.3)	1.00	
Female	122 (89.7)	1.2 (0.49, 2.97)	0.692	268 (85.6)	0.51 (0.23, 1.12)	0.094	91 (71.7)	0.77 (0.36, 1.66)	0.504
Age									
≤20	53 (88.3)	1.64 (0.58, 4.64)	0.349	212 (89.1)	1.51 (0.77, 2.98)	0.229	86 (73.5)	1.58 (0.77, 3.23)	0.214
>20	144 (87.8)	1.00		160 (82.9)	1.00		45 (69.2)	1.00	
Goal commitment (graduation)									
Yes							124 (74.7)	3.77 (1.18, 12.02)	0.025
No							6 (42.9)	1.00	
Environmental Covariates									
Credit load									
Part-time	20 (62.5)	1.00		40 (54.8)	1.00		60 (51.7)	1.00	
Full-time	177 (92.2)	5.11 (1.94, 13.42)	0.001	332 (92.7)	9.44 (4.70, 18.96)	0.000	73 (82.0)	2.93 (1.44, 6.01)	0.003
GPA									
Below 50th percentile	90 (81.1)	1.00		132 (71.4)	1.00				
Above 50th percentile	107 (94.7)	4.10 (1.43, 11.71)	0.009	240 (97.6)	13.89 (5.54, 34.80)	0.000			
Hours spent working									
0 hours per week	79 (91.9)	5.87 (1.28, 26.96)	0.023						
1-10 hours per week	66 (89.2)	4.69 (1.05, 20.91)	0.043						
11-19 hours per week	44 (86.3)	3.70 (0.79, 17.25)	0.096						
20+ hours per week	8 (61.5)	1.00							

Note. OR=odds ratio; CI=confidence interval; GPA=grade point average.

[a] n=224

[b] n=431

[c] n=180

entered into the multivariable model, the elimination process could not be employed. Therefore, the saturated model and the reduced model were the same, and the likelihood ratio test could not be performed and the fit of the model could not be assessed. The results of this model should be interpreted with caution.

The reduced multivariable retention model for students enrolled in parchment programs contained two environmental covariates. Enrolment in full-time studies increased the estimated odds of retention compared to enrolment in part-time studies (OR=9.44, p=0.000). Having a GPA above the 50^{th} percentile also increased the estimated odds of retention (OR=13.89, p=0.000). The results of the likelihood ratio test indicate that the reduced model was not a significant improvement in the fit of the model compared to the saturated model (Chi Square=2.67, p=0.263). In other words, removing the variable for Grade 12 Math from the multivariable model did not significantly improve the fit of the model.

In addition to age and gender, the reduced multivariable model for retention among students enrolled in upgrading and non-parchment programs contained one input and one environmental covariate. Those who felt it was important to graduate had an increase in the estimated odds of retention (OR=3.77, p=0.025). As with the other models, those enrolled in full-time studies had an increase in the estimated odds of retention (OR=4.17, p=0.000) compared to those enrolled in part-time studies. The results of the likelihood ratio test indicate that the reduced model was not a significant improvement in the fit of the model compared to the saturated model (Chi Square=0.48, p=0.786). In other words, removing the variable for Institutional Commitment did not significantly improve the fit of the model.

Comparison of the models

Several similarities and differences existed between the reduced multivariable models for each credential type. Full-time enrolment was a significant predictor of retention in all three models. GPA was also a significant predictor of retention in both the university transfer and parchment models, but not in the upgrading model. The university transfer model was the only model in which the number of hours worked per week remained significant. In addition, the upgrading/non-parchment program was the only model in which goal commitment remained significant.

The greatest differences can be noted between the model for upgrading/other non-parchment programs and the models for university transfer and parchment programs. The results of the model for upgrading/non-parchment programs showed that placing greater importance on graduation increased the estimated odds of retention among students. Goal commitment was associated with retention exclusively among students enrolled in upgrading/non-parchment programs. In addition, having a GPA above the 50^{th} percentile was not significantly associated with retention (after adjusting for other covariates) among students enrolled in upgrading and other non-parchment programs.

Summary of the Results

The overarching goal of this research is to produce information that institutions can use as the basis for improving the success and retention of the students that they admit, register and convocate.

The estimated rate of retention among this sample of Canadian college students was 83.6%. Although the retention rate was similar among students enrolled in the two community

colleges whose data were included in the analysis, the retention rate differed significantly by credential type. Students enrolled in university and transfer programs were significantly more likely than students enrolled in upgrading and non-parchment programs to be retained between the fall and winter semesters.

The data used in these analyses are believed to be suitable for the development of predictive models. The data have adequate representation within each response category, do not exhibit widespread collinearity, have a small proportion of missing values, and have sufficient power within each subset to detect a true effect.

Similarities existed between all of the reduced multivariable models in terms of the covariates that remained significant. The results show that among students in this sample, retention is primarily associated with environmental covariates (with the exception of students enrolled in upgrading/non-parchment programs). In all six of the reduced models, enrolment in full-time studies increased the estimated odds of retention. In five of the six models, having a GPA above the 50^{th} percentile also increased the estimated odds of retention. The remainder of covariates that were significantly associated with retention in the reduced multivariable models were isolated to the respective models. Interpretation of the results, and their implications for stakeholders, will be discussed in the next chapter.

CHAPTER 5: DISCUSSION

An academic institution's ability to ensure that students enrol and re-enrol is central to its function (Astin, 1975; Hagedorn, 2005). The dynamics of student retention have been a topic of interest for nearly a century. Students who attend 2-year colleges may be at greater risk for attrition than those attending 4-year institutions, and the first semester of studies may represent the point at which students are at greatest risk for attrition (Cofer and Somers, 2000). Attrition potentially represents a major loss to the student, to the institution, and to society (Canadian Council on Learning, Oct. 2005).

There is abundant literature identifying factors related to student retention among American student samples (similarities and differences between the current findings and those in the literature will be discussed later in this Chapter). However, a gap exists in the research literature on retention at Canadian community colleges; within the sparse Canadian literature only a few themes exist. Regardless, previous research has indicated that retention is a multi-faceted issue. The determinants of retention may be unique to each context and group examined. The dynamics of retention may depend on the student input characteristics and the environmental context in which students find themselves. Shifting demographics among Canadian college students, and the shifting environment in which colleges operate, suggest that student retention should be re-examined in new ways in order to maintain a current understanding of the issue (Reason, 2009).

This chapter provides a review of the focus of this research and a discussion of the findings. The findings are interpreted within the context of existing literature on Canadian community college student retention, and within the context of the broader literature. The results

will also be considered within the framework of Astin's I-E-O model. The implications of this research for students, institutions, and society are discussed, along with the limitations and strengths of this research. Recommendations for stakeholder that may improve Canadian college student retention will then be provided, as will recommendations for future research.

Summary of Research

The rationale for this research was to produce information that Canadian community colleges can use as a basis for improving the success and the retention of the students that they admit, register and convocate. With looming negative population growth, accountability frameworks, and funding realities, Canadian colleges may not be able to afford to ignore their high rates of attrition in the future. It will become increasingly important to examine institutional retention strategies in order to maintain current enrolment numbers and produce sufficient numbers of graduates to supply labour market demand. This research provides information that may help institutions identify subgroups of students who are most likely to be retained, and develop locally appropriate strategies to improve retention.

Consideration of Findings

The following section provides a discussion of the findings of the analyses related to each research question. The findings are discussed within the context of the literature on Canadian community colleges when available, and in the context of broader literature where appropriate.

Research question 1.a: retention rates.

The first sub-question focused on the estimated rate of retention among Canadian community college students. Estimates in the broad literature for within-year retention at 2-year colleges range from 57% to 83.9% (Cofer and Somers, 2000; Voorhees; 1987; Hippensteel, St.

John, & Starkey, 1996; St. John and Starkey, 1994). Specifically within the sparse literature on Canadian college samples, attrition (rather than retention) rates have been estimated, and these estimates vary between studies. Two Canadian studies examined year-to-year attrition/retention. The SIAST Retention Study (see Sarkar, 1993) produced an estimated attrition rate of 14.5% (which equates to a year-to-year retention rate of 85.5%). More recently, Fisher and Engemann (2009) estimated the year-to-year attrition rate to be 37.7% (which equates to a retention rate of 62.3%). Only one Canadian study estimated within-year attrition. Kriby and Sharpe (2001) found that among students enrolled in an engineering technology program at a Canadian community college, 24.9% were not retained between the fall and winter semesters (which equates to a retention rate of 75.1%).

The estimated overall rate of retention in the current study was 83.6%. This estimate is fairly consistent with previous estimates in the Canadian literature specifically, and the broad literature generally, despite differences in the samples examined and definitions used. Moreover, 83.6% may be a fairly sound estimate of retention at Canadian community colleges, since the estimated retention rate did not differ significantly by institution (83.9% vs. 82.9%) despite the fact that the two institutions had very different student sample characteristics. One would assume that with such different sample characteristics, and with different institutional environments, the retention rate might vary significantly between institutions. The fact that the rates did not differ significantly by institution despite sample and environmental differences lends credibility to the estimates produced by this research. Further, the fact that the data used in this study are generally reflective of the enrolment mix at community colleges increases the likelihood that the results may be generalizable to other institutions.

The estimated retention rate for each credential type varied between 71.9% and 88.0%. This reflects a 16% difference, with university transfer students having the highest retention rate and upgrading/non-parchment students having the lowest retention rate. This was the first study of Canadian college students to examine retention rates by credential; therefore a credential-specific comparison of the results with published literature was not possible. Still, the estimates for retention within each credential type are fairly consistent with published estimates in the literature on broad Canadian college samples (although the rate for university transfer students is slightly higher than published estimates).

Given the consistency among institution-specific estimates, it was surprising how different the estimates were by sub-groups of students enrolled in each credential. The difference between groups indicates that true sub-groups of student may exist by credential type. The difference between groups also alludes to the notion that the factors influencing retention may also differ by credential type. This will be discussed further below.

As previously noted, the current estimates of retention may be slightly inflated since students who withdrew prior to the add/drop deadline were not included in the analysis. This pertained to a relatively small number of cases; it is unlikely that the estimated retention rate would be affected by more than a few percentage points. This may have been avoided if accommodations were made to include these students in the analysis; this will be discussed below in the section on Recommendations.

Research question 1.b: suitability of the data for a predictive model

The second sub-question focused on the suitability of the data for the development of a predictive model. Doing so ensured that the analysis could proceed as planned, and that the

results would be meaningful to the institutions that may potentially use the information produced from this study. The appropriateness of the data were evaluated using four methods that together provide a comprehensive understanding of the data: examining the distributions of variables, examining inter-relationships among variables, evaluating the percentage of missing values in each variable, and conducting power analyses. Based on the results, the data are believed to be highly appropriate for development of predictive models. The data generally have adequate representation within each response category, do not exhibit widespread collinearity, have a small proportion of missing values, and have sufficient power within each subset to detect a true effect in regression models.

Since none of the published Canadian studies on college student retention used logistic regression, published studies in the broader literature were referenced for comparison. Appendix D provides an overview of published studies that examined retention using logistic regression. Comparability of the appropriateness of the data used in the current study to those in Appendix D is difficult since the same level of detail is not provided in the majority of published literature. However, some other sources allude to a similar process of examination of data distributions with cross-tabulations using Chi Square tests for association (Nguyen et al, 2010, Fike and Fike, 2008), examination of missing values, and consideration for potential collinearity (Fike and Fike, 2008). Since similar methods were used in the current research, it is likely that the appropriateness of the data were properly evaluated.

Sample size is the only parameter consistently discussed across studies using logistic regression. Among the studies listed in Appendix D, the sample sizes ranged from 310 to 9,200, with the majority of studies using sample sizes well over 1000. The sample sizes for the models

in the current study ranged from 180 to 837. While the sample sizes used in the models in the current study are substantially smaller than those generally used in the published literature using logistic regression, the power analysis for each model showed adequate power for the analyses. Therefore for comparability purposes, the sample sizes used within the current study are considered adequate (although generally smaller than the sample sizes published research).

The inclusiveness of the sample used in this research also increases the appropriateness of this data for the analysis, even though this was not overtly evaluated. Hagedorn (2005) recommended that the data used to develop predictive models of retention be reflective of the true situation facing community colleges. Perhaps one of the greatest strengths of this research, which made this data appropriate for the development of statistical models, was the fact that it is derived from a sample that is representative of the enrolment profile at many community colleges. Most previous research has used very limited samples when developing predictive models of retention, which may potentially produce misleading results (Pittman, 2008). While the majority of research using logistic regression has limited inclusion criteria to those who are first-time and/or full-time students (Settle, 2011; Fike & Fike, 2008; Nguyen et al., 2010; Bers, 1988; Wells, 2008), inclusion in the current study was open to the entire cohort of students enrolled in the first year of credit programming (with the exception of apprenticeships, ESL students, and short-duration credentials) at the two colleges involved. Therefore students who may be at higher risk for attrition, such as part-time students, were included in this analysis. Compared to previous published research on predictive models of retention, the results of the current study may produce a picture of retention that is more representative of the true situation facing community colleges.

Research question 1.c: factors predicting within-year retention

The dynamics of retention are constantly evolving and depend on the unique characteristics of each particular cohort of students at each institution. This necessitates that student retention should be re-examined in new ways in order to maintain a current understanding of the issue (Reason, 2009). Yale (2010) recommends that post-secondary institutions use predictive models of student outcomes to inform decision making when possible. Since one of the goals of this research was to produce information that could be used by stakeholders, understanding the dynamics of retention among Canadian college students using a predictive model was an important component of this research. Therefore, the third sub-question in this study asked: what student background and academic characteristics predict within-year retention among those attending two Canadian comprehensive community colleges? To answer this question, a predictive model of retention using the entire sample was developed that considered a broad array of student input and environmental factors. Here, the results will be compared primarily to literature on Canadian college students, since the dynamics of retention for this population may be specific. The broader literature is also referenced to comment on congruencies and discrepancies in the results.

Although the results of previous research on retention among Canadian college student samples are not consistent, certain themes exist. These include associations between retention and past academic performance, student engagement/integration/goal commitment, and certain demographic factors – in particular age and gender (Fisher and Engemann, 2009; Sarkar, 1993; Kirby and Sharpe, 2001; Jorgensen, et al., 2009). Beyond these themes, other factors have also been inconsistently identified, including date of admission, program choice (Fisher and

Engemann, 2009), credit load, occupational uncertainty (Kirby and Sharpe, 2001), use of student support services, disability, aboriginal ancestry, marital status, dependent children, employment, lower educational goals, and understanding of career choices (Sarkar, 1993). These themes are generally reinforced by the broader literature, although due to the abundance of studies in the broader literature a number of additional associations have also been identified (see Appendix A).

Numerous methodological differences complicate the comparison of published results on Canadian college samples to those of the current study. Differences in samples used, methods employed, and variables considered cloud the results. In addition, the various definitions of retention/attrition used further complicate comparison of results. Published research on Canadian college student retention actually tends to address attrition, as opposed to retention. The current study took the opposite approach and attempted to identify factors predicting retention, as opposed to attrition. This distinction is an important consideration when comparing the results of the current study to those of past research.

Among the aggregate sample used in the current study, retention was predicted by full-time enrolment and having a GPA above the 50th percentile (after adjusting for age and gender). While these results confirm some of the findings of previous research on retention among Canadian college students, they are also divergent in a number of ways.

Only one other published Canadian study examined credit load as an independent variable. Kirby and Sharpe (2001) found that Canadian college students enrolled in full-time studies were more likely to be retained than those enrolled in part-time studies. Kirby and Sharpe state that a relatively small number of students included in their study were enrolled part-time, so

the results should be interpreted with caution. The results of the current study support those of Kirby and Sharpe with regards to credit load: full-time enrolment was a strong predictor of retention in the current study. Given that a substantial proportion of students included in the current study were enrolled part-time (about 24%), the results suggest that credit load may in fact be a significant predictor of retention among Canadian college students. Moreover, the results of several studies in the broader literature also suggest a link between credit load and retention, further supporting the findings of the research (Fike and Fike, 2008; Cofer and Somers, 2010; Settle, 2011; Herzog, 2005).

The association between first-semester GPA and retention has not been examined in published Canadian college literature, but has been widely confirmed in the broader literature (Settle, 2011; Grayson, 1998, Cofer and Somers, 2010; Ngyuen et al, 2010; Liu, 2010; Braunstein, McGrath, and Pescatrice, 2000/2001; Reason, 2009; Herzog, 2005; Ma and Frempong, 2008). In line with these findings, the results of the current study identified first-semester GPA as a predictor of retention in both univariable and multivariable analyses. Students who had a GPA above the 50^{th} percentile were over four times more likely than those with a lower GPA to be retained.

Age has been shown to be associated with retention in the majority of Canadian literature on college samples (Fisher and Engemann, 2009; Sarkar, 1993; Jorgensen, et al., 2009). In the Canadian literature, greater age has shown to be inversely associated with retention, so that older students tended to have a lower rate of retention compared to younger students. These results for the effect of age on retention are mixed in the broader literature (Ngyuen et al., 2010; Settle, 2011; Fike and Fike, 2008; Cofer and Somers, 2010; Bers, 1988; Robertson and Taylor, 2009;

Reason, 2009). Differences in the samples used and methods employed may produce mixed results.

The results showing the effect of gender on retention among the Canadian college student studies were mixed. Some results suggest that males (Fisher and Engemann, 2009) – particularly males with a lower high school average – are more likely than females to drop-out (Jorgensen et al., 2009). Conversely, female gender has also been associated with a higher rate of attrition (Sarkar, 1993). The broader literature also suggests mixed results regarding the association between gender and retention (Grayson, 1998; Bers, 1988; Reason, 2009; Ma and Frempong, 2008).

Because the associations between retention and age, and retention and gender, have been noted in the literature on Canadian college samples, they were included in the univariable and multivariable analyses in the current study (this approach has also been used in published research using similar methods – see Fike and Fike, 2008). However, contrary to previous research on Canadian college students neither factor was a highly significant predictor of retention in univariable or multivariable analyses. These results suggest that among this sample of Canadian college students, age and gender may not predict retention. This reinforces the idea that the dynamics of retention may be unique to the cohort of student being studies, necessitating that context-specific examinations of retention be conducted.

Published literature on Canadian college students (Fisher and Engemann, 2009; Sarkar, 1993; Kirby and Sharpe, 2001; Jorgensen, et al., 2009), and many studies in the broader literature (Fike and Fike, 2008; Ngyuen et al., 2010; Herzog, 2003; Marshall, 2008) have suggested that measures of past academic performance may be associated with retention. That being said,

among these studies past academic performance has been defined in various ways – for example in terms of high school GPA, performance in a grade 12 math course, and as a more general measure of academic preparedness or ability. In the current study, two measures of past academic performance were included: high school performance and whether or not a student passed grade 12 pure math. Better high school performance was associated with increased odds of retention in univariable analysis of the aggregate data, but did not remain a significant predictor in multivariable analysis. Passing grade 12 math was not associated with retention in univariable or multivariable analyses.

Given that previously published research on student retention has indicated the significance of past academic performance, it is surprising that the two measures of past academic performance used in the current study (high school performance, pass/fail grade 12 pure math) were not more strongly associated with retention. In the current study, it is possible that some forms of bias may have skewed the results. High school performance and passing/failing grade 12 pure math were both self-reported measures; students may have tended to portray themselves in a more socially desirable manner (social desirability bias), or may have biased recollection of past events (recall bias), thereby inflating estimates of their past academic performance. Supporting this notion is the fact that older students tended to be more likely than younger students to report a better high school average (see Appendix E). Potentially, a longer duration of time between high school attendance and enrolment in college may affect one's memory of high school performance. Ideally this research would have collected information on past academic performance from student records rather than via self-report; however it was not routinely available from student records.

This research also examined the association between a number of other factors identified in the literature on retention. After adjusting for other covariates, the majority of these factors were not found to significantly predict retention in this analysis. These factors include parental education (Fike & Fike, 2008; Nguyen et al., 2010; Wells, 2010; Hand & Payne, 2008; Grayson, 1998; Ishitani, 2003), marital status (Reason, 2009; Sarkar, 1993; Ma & Frempong, 2008), academic aspirations (Cofer & Somers, 2010; Sarkar, 1993; Ma & Frempong, 2008), time spent studying while in college (Madgett & Belanger, 2008), having friends attend the same institution (Settle, 2011; Roberts & Styron, 2010; Wells, 2008), having family attend the same institution (Settle, 2011), family emotional support (Robertson & Taylor, 2009), living arrangements (Settle, 2011; Madgett & Belanger, 2008; Liu, 2010; Herzog, 2005; Grayson, 1998), time spent working while in college (Grayson, 1998; Sarkar, 1993), and financial assistance (Settle, 2011; Fike & Fike, 2008; Cofer & Somers, 1998, 2010; Madgett & Belanger, 2008; Herzog, 2003; McElroy, 2005; Ma & Frempong, 2008), among others. While the divergence of these results may be partially attributable to differences in the samples used and methods employed, they may also reinforce the idea that retention is context-specific. The determinants of retention may depend on the unique characteristics of a given cohort of students and the unique characteristics of the institutional environment.

Research question 1.d: predictors of retention by institution

Although the third sub-question sought to understand the dynamics of retention among the total sample used in this research, it was still unknown whether or not the dynamics differed by college. The fourth sub-question asked: Do the predictive models for retention differ between these two institutions?

This is the first Canadian study to examine and compare predictors of retention among more than one community college. Published studies of Canadian college students have focused on samples from single institutions (occasionally over multiple yeas). In the broader literature, analyses stemming from the Canadian *Youth in Transition* survey data have examined retention among post-secondary students in general (as opposed to college and university students separately), and examine system-level retention rather than retention at individual institutions (Parkin and Baldwin, 2009). In the American context, analyses of the *National Postsecondary Student Aid Survey* data have examined retention among two-year college students in general as opposed to separate institutions (Cofer and Somers, 2000; Hippentsteel et al, 1996; St. John and Starkey, 1994). The results of the current study add to our understanding of retention among Canadian college students.

Given that the retention rates at the two institutions studied were similar, and that institution of attendance did not significantly contribute to retention in the aggregate data, the similarities and differences between the two institutional models were interesting. Retention at both College A and B was predicted by credit load and GPA. Gender also predicted retention at College A, but not at B. The results suggest that, among the input and environmental factors for which data were available, those that predict within-year retention among students attending these two Canadian colleges may be similar. However, within the overall similarities distinct differences existed.

The results suggest that retention at both colleges was most strongly predicted by full-time credit load and having a GPA above the 50th percentile. Students who were enrolled full-time were 5.3 times more likely to be retained at College A, and 19.2 times more likely to be

retained at College B compared to students enrolled part-time. Students with a GPA above the 50th percentile were over five times more likely to be retained at College A and 18.6 times more likely to be retained at College B compared to students with a GPA below the 50th percentile.

Although direct comparison of these results to those found in the literature is not possible, the broad literature generally supports the finding that credit load and GPA are associated with retention. Multiple studies – both within the Canadian context and in the broader literature – suggest an association between credit load and retention (Kirby and Sharpe, 2001; Fike and Fike, 2008; Cofer and Somers, 2010; Settle, 2011; Herzog, 2005). Similarly, the broader literature has consistently identified a relationship between GPA and retention (Settle, 2011; Grayson, 1998, Cofer and Somers, 2010; Ngyuen et al, 2010; Liu, 2010; Braunstein, McGrath, and Pescatrice, 2000/2001; Reason, 2009; Herzog, 2005; Ma and Frempong, 2008). Therefore it is not surprising that credit load and GPA were the strongest predictors of retention among both College A and B.

Despite having similar predictors of retention, the magnitudes of the effects observed for these predictors were widely different between institutions. For example, consider the ORs for GPA at College A and B (mentioned above). There is over a 13 point difference in the magnitude of the ORs for GPA between institutions. Similar results can be seen for the effect of credit load. The differences in the relative magnitude of factors contributing to retention noted between the two institutional models suggest that the dynamics of retention should be examined on an institution-by-institution basis.

These differences in the magnitude of effects could be partly attributable to a smaller sample size at College B compared to College A. While the reduced model for College A

contained data from 644 students, the reduced model for College B contained data from only 193 students. A small sample size can increase the magnitude of the ORs and the decrease the precision of the estimates. However, the model for College B had sufficient power, and the sample size should have theoretically allowed for up to 19 covariates (ten observations per covariate).

At the same time, the observed difference may represent a true effect given that the data distributions in Table 5 show that significant differences existed between College A and B in terms of GPA and credit load. Compared to College A, College B had a significantly smaller proportion of students who had a GPA above the 50^{th} percentile, and who were enrolled full-time. This supports the notion that although the same drivers may be influencing retention at both institutions, they may be doing so to different extents in different contexts.

The results regarding the effect of gender on retention at College A are interesting. In the reduced model for College A, females had lower odds of being retained compared to males. In cross-tabulations, the retention rate did not differ significantly by gender. Gender was also not a significant predictor of retention in univariable analysis for either institution, yet became a significant predictor of retention among students enrolled at College A in the reduced multivariable analysis.

Although evaluating the potential for confounding is beyond the scope of this analysis, it is plausible that interactions between gender, GPA, and retention may have produced a spurious association between gender and retention. Gender and GPA were significantly associated with each other in addition to being associated with retention (data not shown). Females at College A

were significantly more likely than males to have a GPA above the 50th percentile. This suggests that the observed effect may be due to a spurious association between covariates.

The results of previous research on student retention have been mixed with regards to the effect of gender and therefore do not aid in the interpretation of the results (Grayson, 1998; Bers, 1988; Reason, 2009; Ma and Frempong, 2008). Within the Canadian literature on college student samples, Sarkar (1993) found that females were less likely than males to be retained, but Fisher and Engemann (2009) and Jorgensen et al. (2009) found the opposite results. The lack of consistency in the results of previous research, and the mixing of effects suspected in the current data suggest that the association between gender and retention observed in this study may be spurious.

As with the aggregate model in sub-question 1.c, the potential effects of a number of factors suggested to be associated with retention in the literature were examined in this analysis (see above for a list of factors). However, after adjusting for other covariates, the majority of these factors were not found to significantly predict retention in at College A or B. While the incongruence of these results may be partially attributable to differences in the samples used and methods employed, they again reinforce the idea that retention is context-specific. The determinants of retention may be specific to unique cohort and environmental characteristics at each isntitution.

Research question 1.e: predictive models by credential type

Yale (2010) recommends that student data be examined by sub-groups of the total sample. Although the predictors of retention identified among the aggregate sample and each institution were similar, it was unknown whether or not the same predictors would be identified

among sub-groups of students by credential type. Sub-question 1.e asked: do the predictive models differ for students enrolled in university transfer programs, upgrading and other non-credential programs, and parchment programs leading to a credential? To answer this question, three predictive models were developed and compared to examine the dynamics of retention among each credential.

There is a paucity of published literature examining retention among sub-groups of students who together represent the enrolment mix at most colleges. Previously published studies of retention among Canadian college students have not examined retention among their samples by sub-group, nor has the majority of the research in the broader literature. Settle (2011) developed predictive models of retention for first-generation students and continuing generation students, and found that significant differences existed between groups. Crawford Sorey and Harris Duggan (2008) demonstrated that the predictors of retention differed among traditional-aged students and adult students. Bers (1988) found that different factors were associated with retention by major. Similar to these studies, the current study found that examining the predictors of retention by sub-groups yielded interesting and potentially useful results.

The results of the current research suggest that different predictors of retention exist for certain sub-groups of Canadian college students by credential type. Specifically, university transfer students, and those enrolled in upgrading/other non-parchment programs, may represent sub-groups of students for whom the dynamics of retention are unique.

As with the aggregate model, retention among students enrolled in university transfer programs was predicted by credit load and GPA. However, retention was also predicted by the amount of time students spent working per week. University transfer students who did not work,

and those who worked ten or fewer hours per week, were about 5.9 and 4.7 times (respectively) more likely to be retained than those who worked more than 20 hours per week. Sarkar (1993) also found that students (in general) who were employed and worked a greater number of hours per week were less likely to be retained. The results of the current research are the first to suggest that the link between employment and retention exists specifically among university transfer students. However, these results should be interpreted with caution since the fit of the model could not be evaluated.

Students enrolled in university transfer programs may represent a distinct group among whom the dynamics of retention are unique. They were the only group for whom the number of hours spent working was a significant predictor of retention. Furthermore, the profile of university transfer student characteristics presented in Table 5 shows that in many ways they are distinct from students enrolled in parchment and upgrading/non-parchment programs. The characteristics and needs of university transfer students may require greater consideration in the future.

Among student enrolled in upgrading and other non-parchment programs, retention was predicted only by a student's commitment to the goal of graduation and credit load. Students who expressed placing greater importance on the goal of graduation were nearly four times more likely than those who did not express the same commitment to be retained. Goal commitment was also found to be associated with retention among college students in general in the Canadian literature (rather than specifically among students enrolled in upgrading/non-parchment programs) (Sarkar, 1993).

As with the other models, credit load predicted retention among students enrolled in upgrading/non-parchment programs. Those enrolled full-time were nearly three times more likely to be retained (OR=2.93) than those enrolled on a part-time basis. This finding is generally supported by literature among general student samples. In addition, students enrolled in upgrading and other non-parchment programs were the only sub-group examined for whom retention was *not* associated with GPA. GPA was not associated with retention in univariable or multivariable analyses among this sub-group, indicating that academic performance may not be an important factor contributing the retention among students enrolled in upgrading/non-parchment programs. This finding is incongruent with findings in the literature on general student samples (mentioned above), which suggest links between GPA and retention. This finding represents a unique contribution of the current study, since it is one of the first to examine retention by credential. Further, the difference noted between students enrolled in upgrading/non-parchment programs, and those enrolled in the other credential types examined suggest that unique factors contribute to retention among this group of students.

As with the aggregate model, credit load and GPA were found to predict retention among students enrolled in parchment programs. Having a GPA above the 50th percentile and full-time enrolment were the only covariates that remained significant in the reduced multivariable model. Students enrolled in parchment programs who were enrolled full-time were nearly 9.5 times more likely than students enrolled part-time to be retained. Those who had a GPA above the 50th percentile were nearly 14 times more likely than those with a lower GPA to be retained. The predictors of retention among this sub-group of students are relatively 'conventional' compared to the aggregate results, and compared to the results of analyses for other sub-groups. The main

difference lies in the magnitude of the estimates for GPA and credit load, which were far higher than those found in the other models (GPA OR=13.89, full-time enrolment OR=9.44). This suggests that retention among students enrolled in parchment programs are quite strongly predicted by these two variables.

The differences in the relative magnitude of common factors contributing to retention noted between the three credential-types suggest that the dynamics of retention should be examined by sub-group. Although the predictors of retention in these samples were similar, their estimated effects at each credential were quite different. For example, students enrolled in upgrading/non-parchment programs who were enrolled full-time were nearly three times more likely than those enrolled part-time to be retained. University transfer students enrolled full-time were over five times more likely than those enrolled part-time to be retained, and students enrolled full-time in parchment programs were almost 9.5 times more likely that part-time students to be retained. Similar differences were observed of the effect of GPA on retention among students enrolled in university transfer and parchment programs. Whereas university transfer students with a GPA above the 50th percentile were just over four times as likely to be retained as those with a GPA below the 50th percentile, students enrolled in parchment programs with a GPA above the 50th percentile were nearly 14 times as likely to be retained compared to those with a GPA below the 50th percentile.

Summary of results

Relatively few published studies have examined determinants of retention among Canadian students in general, let alone among Canadian community college students. The results of the current study add to the sparse literature on retention among Canadian college students.

The results estimate that the retention rate among Canadian college students is comparable to published estimates. The fact that the retention rate did not differ by institution, despite having significantly different student and environmental characteristics, lends credibility to the results. Differences in the retention rate may exist by credential type; this difference has not previously been recognized in the literature.

The results of the multivariable models predicting retention also provide valuable information. The models can be compared in Table 20, which provides a visual representation of the covariates that predict retention in each of the six models. The directionality of significant relationships between retention and each covariate are represented as 'positive' (increased odds of retention) or 'negative' (decreased odds of retention) in each of the reduced models. The most obvious trend in Table 20 is the consistent positive effect of full-time enrolment and GPA above the 50^{th} percentile on retention. The greatest differences can be noted between the multivariable models for each credential.

Table 19

Summary of Significant Associations with Retention in all Reduced Multivariable Models

	Model					
		Institutional		Credential		
Covariates	**Aggregate**	**College A**	**College B**	**University Transfer**	**Parchment Programs**	**Upgrading & Non-Parchment Programs**
Student Input Covariates						
Gender						
Male						
Female		Negative				
Goal commitment (graduation)						
Yes						Positive
No						
Environmental Covariates						
Credit load						
Part-time						
Full-time	Positive	Positive	Positive	Positive	Positive	Positive
GPA						
Below 50th percentile						
Above 50th percentile	Positive	Positive	Positive	Positive	Positive	
Hours spent working						
0 hours per week				Positive		
1-10 hours per week				Positive		
11-19 hours per week						
20+ hours per week						

Note. GPA=grade point average.

The results of the current study were consistent with some of the determinants of retention noted in existing literature on Canadian college students, but were also divergent in several ways. In addition, the results of the current study have suggested that, among Canadian college students, sub-groups may exist for whom retention is predicted by unique factors. When considered as a whole, the results of this research suggest some interesting trends.

Credit Load

The results of this research reinforce those of Kirby and Sharpe (2001) and others in the broader literature who found a possible association between full-time enrolment and retention. The results of the current study suggest that credit load may be one of the most important predictors of retention among this sample. Credit load was associated with retention in all six multivariable models developed in this research.

GPA

The combined results of this research indicate that academic performance in the first semester of studies (as measured by GPA) is generally an important predictor of retention. GPA was a significant predictor of retention in five of the six multivariable models developed. These results are consistent with previous research, in which GPA has been shown to be associated with retention.

Age and gender

Age and gender were also shown to be associated with retention in the majority of Canadian literature on college samples. Contrary to previous research neither factor was a highly significant predictor of retention in any of the analyses. Gender became a significant predictor in only one model, and this was likely due to a spurious association. Based on these results, there is little evidence to suggest that age and gender are predictors of retention among this sample.

Academic performance

Although previous research on Canadian college student samples has indicated a link between retention and past academic performance, the link was not confirmed by the measures of past academic performance used in the current study. Although some associations existed between measures of past academic performance and retention in univariable analyses, they did not remain significant in multivariable analyses once other factors were considered. Although some forms of bias may have influenced the data, the results of this research do not suggest that past academic performance is as a predictor of retention among this sample.

Predictors of retention among sub-groups

The results of the current research suggest that different predictors of retention exist for certain sub-groups of Canadian college students. In particular, the dynamics of retention among students enrolled in university transfer programs and upgrading/non-parchment programs may be unique. Retention among these sub-groups should be examined separately.

Consideration of Results within Methodological Framework

This research was conducted within the framework of Astin's I-E-O model (1991; 1984). Astin's model states that institutional outputs such as retention must be evaluated in the context of student input and environmental factors. Factors that are inherent to the student may influence the ways in which the student reacts and interacts with environmental factors found within each institutional context.

A broad range of student input and environmental factors were considered for inclusion in this research. Consistent with Astin's theory, several of the student input and environmental factors were significantly associated with each other (see Appendix E). It is plausible that many

of the student input factors interacted with the environmental factors to influence the likelihood of retention.

Potential interaction between student input and environmental factors in the models can be observed by comparing the change in the point estimates of covariates between the univariable and multivariable models. Generally, a change in the point estimates greater than 0.10 between the univariable and multivariable regression models may suggest that an interaction between covariates may be occurring. For example, in the results for Research Question 1.c, the univariable OR for retention among students enrolled full-time was 6.92 (see Table 10). In the multivariable analysis, the OR for retention among full-time enrolment had decreased to 5.92 (see Table 11). Since the multivariable model adjusts the point estimates for the effect of other covariates in the model, this change may suggests that other covariates in the model influenced the odds of retention among full-time students. It is beyond the scope of this research to speculate at the cause of change in each point estimate. However, observing these interactions may suggest areas where student input and environmental variables are interacting.

More useful information may be gleaned from observing the overall trends between models with regards to the student input and environmental factors that predicted retention. Although many student input factors were considered for inclusion in this research, most failed to remain significant through the elimination process used in the multivariable analysis. The exception to this was the significance of goal commitment in the multivariable model for upgrading/non-parchment students.

Taken as a whole, the results of the multivariable models developed as part of this research suggest that it is mainly environmental factors that predict retention among this sample. The results indicate that certain environmental factors may play a greater role in the dynamics of

Canadian college student retention than input factors. In most cases, certain aspects of the educational environment may be more directly associated with retention than characteristics inherent to the student. As shown above, it is also important to be cognizant of the potential interaction between student input and environmental factors.

Two environmental factors - credit load and GPA - were consistently strong predictors of retention. Most prominent among these environmental factors was credit load. Compared to those enrolled in part-time studies, those enrolled in full-time studies had increased odds of retention in all six multivariable models. Within four of the six multivariable models, the magnitude of the point estimates for full-time enrolment were greater than those for other covariates. Within the other two multivariable models (those for parchment programs and upgrading/non-parchment programs), full-time enrolment was still a highly significant predictor of retention. GPA was also a prominent environmental factor that was associated with retention in the results of this research. GPA was a strong predictor of retention in five of the six multivariable models (the exception of the retention model for students enrolled in upgrading/non-parchment programs).

It is surprising that other environmental factors did not also predict retention in this research. Environmental factors showing an association with retention in the literature on Canadian college student samples include student engagement/integration (Fisher and Engemann, 2009; Kirby and Sharpe, 2001; Jorgensen, et al., 2009), program choice (Kirby and Sharpe, 2001), use of student support services, employment, and understanding of career choices (Sarkar, 1993). The role of majority of these factors could not be assessed in the current study. A number of environmental factors shown to be associated with retention in the broader literature were included in this study, but did not prove to predict retention among this sample. These

included time spent studying while in college (Madgett & Belanger, 2008), having friends attend the same institution (Settle, 2011; Roberts & Styron, 2010; Wells, 2008), having family attend the same institution (Settle, 2011), family emotional support (Robertson & Taylor, 2009), living arrangements (Settle, 2011; Madgett & Belanger, 2008; Liu, 2010; Herzog, 2005; Grayson, 1998), time spent working while in college (Grayson, 1998; Sarkar, 1993), and financial assistance (Settle, 2011; Fike & Fike, 2008; Cofer & Somers, 1998, 2010; Madgett & Belanger, 2008; Herzog, 2003; McElroy, 2005; Ma & Frempong, 2008), among others. The fact that these factors were not found to predict retention in the current study reinforces the notion that the dynamics of retention may differ by cohort and institutional environment. Retention needs to be examined on a contextual basis.

Only one student input factor remained significant in one of the six multivariable models. Expressed goal commitment predicted retention among students enrolled in upgrading/non-parchment programs. It was surprising that this was the only input factor found to predict retention in this research – and only among a specific sub-group of students at that. Previous research has shown numerous input factors to predict retention among college student samples. Within the Canadian literature on college student retention, multiple student input factors have been associated with retention, including age, gender (Fisher and Engemann, 2009; Sarkar, 1993; Kirby and Sharpe, 2001; Jorgensen, et al., 2009), occupational uncertainty (Kirby and Sharpe, 2001), disability, aboriginal ancestry, marital status, dependent children, and lower educational goals (Sarkar, 1993). A number of these factors were not included in the current research. A number of student input factors shown to be associated with retention in the broader literature were also included in this research, but did not prove to predict retention among this sample. These factors include parental education (Fike & Fike, 2008; Nguyen et al., 2010; Wells, 2010;

Hand & Payne, 2008; Grayson, 1998; Ishitani, 2003), marital status (Reason, 2009; Sarkar, 1993; Ma & Frempong, 2008), and academic aspirations (Cofer & Somers, 2010; Sarkar, 1993; Ma & Frempong, 2008).

Limitations

There are several limitations to this research that need to be considered when interpreting the results. These limitations relate to the single cohort from whom data were collected, the coding of the data, the nature of the variables examined, and the potential for bias to have influenced the results.

Data from a single cohort

The analysis for this research was based on a single cohort of students beginning their studies at two community colleges in western Canada in the fall of 2012. Therefore, the results may not be generalizable to cohorts of students with similar or dissimilar characteristics at other institutions. The results should be interpreted with caution in dissimilar situations.

Coding

The results of this study are also limited because no differentiation/distinction could be made between students who discontinue their studies and those who left to study at another institution (system-level retention). Further, the results do not differentiate between those who dropped out permanently from those students who stopped out temporarily. Systematic differences in the predictors of retention may exist between students who drop-out, stop-out, and fail-out. Data on reasons for leaving are lacking on the majority of students in this sample who left their studies prematurely.

This research may have also overestimated the rate of retention, since students who are enrolled in the fall semester may drop-out before the add/drop deadline. Student who dropped-

out prior to the add/drop deadline at each institution were not included in this analysis. For these students, complete data on several key factors (e.g. GPA, credit load) was not available. However, the number of students who drop-out completely before the add/drop deadline is typically small (<20 per institution), so the effect of their exclusion was likely minimal.

Nature of variables examined

Categorization of continuous variables (such as age or GPA) may mute the effects of these variables on retention. In other words, the observed strength of the relationship between variables may be stronger (or weaker) if the variable is left in its continuous state. However, given the dichotomous nature of the outcome (i.e. retention or attrition) and the complexity of the models, it was felt that the prudent choice would be to dichotomize continuous variables.

It is possible that some variables associated with retention were not included in this research. The variables selected for inclusion were based on characteristics identified from the results of the previous research on the retention of Canadian and American college and university samples cited earlier. Selection of the variables for inclusion in this research was also limited by the availability of data at the two colleges involved. It is possible that some unknown factors may also influence retention among students at these Canadian community colleges. In other cases, data may not have been available or feasible to collect for certain factors previously shown to be influential (e.g. engagement, placement test scores).

Potential for bias

Several forms of bias may have influenced the results. The data were based on survey data that was voluntarily provided to the institutions that collected the data. Students were offered a chance to win a $50 bookstore gift card as an incentive for completing the survey. Therefore, students in financial need may have been more likely to participate. However, the risk

of self-selection bias may have been reduced by the fact that students were not aware of the incentive until the end of the survey. Self-selection bias may have also been introduced by systematic difference between those who chose to supply identifying information and those who did not, or those who attended orientation and those who did not. Due to the self-reported nature of many of the variables, recall bias and social desirability bias may have also influenced the results (Rothman, 2002).

Strengths

This research had several strengths, including the sample used, the design, the broad range of independent variables considered, the definition of retention used, and the examination of retention among sub-groups of students. In many instances, this research represents a unique contribution to the literature on retention among Canadian college students.

Sample

That data used for this research was obtained from a sample of students enrolled in the first year of a program at two community colleges in western Canada. This is one of the first studies to provide a comprehensive demographic and academic profile of Canadian community college students. Unlike most research on retention, this sample reflects the enrolment mix of most community colleges and therefore increases the likelihood that the results may be generalizable beyond the current study.

Design

This was one of the first studies to examine Canadian college student retention using a predictive model. While this design should not be considered superior to other methods, it does allow for causal inferences to be drawn regarding the effect of the factors examined on retention. In addition, the results of this research provide a new perspective on Canadian college student

retention that does not currently exist in published literature. The perspective added by the results of this research supplements the information already in existence and provides a stronger foundation upon which future research may be conducted.

Student input and environmental factors

This research included information on a broad range of student input and environmental factors. This allowed for a broad consideration of factors when building the multivariable models for retention. This increases the likelihood that the variables that remained significant in the multivariable models are in fact predictive of retention among this sample of Canadian college students.

Definition of retention

This research is one of the first studies to examine within-year retention among Canadian college students. The majority of published research has examined year-to-year attrition, which effectively excludes many students enrolled at community colleges. The definition of retention used in the current study allows for inclusion of students enrolled in upgrading/non-parchment, certificate, diploma, university transfer, and applied degree programs. Further, since college students may be at greatest risk for attrition between the first and second semester of studies, this research examines the dynamics of retention during a critical period of time.

Examination of sub-groups

This is one of the first published studies of Canadian college students to examine the dynamics of retention among sub-groups of students. The results contribute to our understanding of the complexities of retention, and emphasize the need for individualized or context-specific approaches to be taken when considering ways to increase retention.

Implications

As previously discussed, students, college personnel, and society have a vested interest in student retention. Information on retention rates and the dynamics of retention may have different implications for each stakeholder group. This section will discuss the potential implications of the results of the current study for each stakeholder group.

Students

Students have a vested interest in their own college completion, since they are the direct beneficiaries of post-secondary education (Canadian Council on Learning, 2009a). Estimates of within-year retention among first-year students indicate that the first semester may be a critical period in a student's academic career (Cofer and Somers, 2000). The results of this research provide prospective college students with a realistic estimate of the likelihood of retention between the first and second semesters. Understanding the dynamics of retention may help students make choices about their education, give them a realistic idea of the potential risks and rewards of attendance (Astin, 1975, 1993; Astin & Antonio, 2012), and identify measures to improve the likelihood of a positive experience while achieving their goals.

The results of the current research may therefore have value for students who are considering attending a Canadian college. The results provide prospective students with a sense of the likelihood of retention at a community college in western Canada. The results also provide prospective students with insight into the dynamics of retention at two colleges in western Canada, and the dynamics of retention by credential type. Prospective students may use the recommendations of this research to help shape their experience within the academic environment.

Community Colleges

College administration and staff has a vested interest in retention rates and the factors contributing to retention among their students. The results of the current study have several implications for staff at community colleges. Information on retention can be used to inform SEM policy and practices and provide realistic estimates of the potential for impact (Astin, 1975, 1993; Astin & Antonio, 2012). Retention rates may also be used as a performance indicator in post-secondary education; information on retention rates can help institutions demonstrate quality and monitor progress toward goals (Canadian Council on Learning, 2009a). Related to this, retention rates can be used to demonstrate return on investment of public funds and alignment with public policy objectives (Bogue & Dandridge Johnson, 2010). Each of these will be discussed in greater detail below.

Retention is an important indicator to institutions because it is often used as part of SEM initiatives. The information produced from this research may help to inform focused and integrated strategies relating to enrolment goals. Implementation of retention strategies based on the information provide by this research may relate to the everyday activities of staff as they interact with students and each other. Institutions may use SEM retention strategies to improve the quality of the educational experience and maintain/increase enrolment. By using data on the dynamics of retention to understand the student experience, institutions can guide their practices to better align with student needs, which may ultimately increase retention (Henderson, 2012).

The results of this research have several implications for institutional policy and SEM strategies aimed at increasing retention. Ensuring staff have a thorough understanding of the factors associates with retention may help institutions develop sounds policies and practices to encourage retention. The results of this research suggest that the greatest gains in retention may

be via SEM strategies aimed at environmental factors - specifically, encouraging full-time enrolment and supporting students' academic performance. The results also suggest several areas where targeted strategies may increase retention among specific sub-groups of students. In addition, the results suggest that certain factors may have a stronger influence on retention than others. This information may be used to inform strategic resource allocation decisions to be made in times of fiscal constraint.

Institutional approaches to enrolment management often default to the development of strategies aimed at increasing retention rates to the detriment of student access (Green, 2013). This most often takes the form of higher admissions standards (Hossler & Kalsbeek, 2013). The results of the current study suggest that SEM strategies such as higher admissions standards may not be the most effective approach to increase retention among Canadian college students, since past academic performance did not remain in any of the final multivariable models. SEM strategies aimed at supporting academic success among current students (thereby maintaining access) may be more effective.

Retention rates may also be used as a performance indicator to help institutions demonstrate quality and monitor progress toward goals (Canadian Council on Learning, 2009a). Although many indicators of quality exist and are used throughout the PSE system, retention tends to be one of the most frequently used quality assurance indicators among institutions (Bogue & Dandridge Johnson, 2010; Tagg, 2008). Retention is an attractive indicator because the definition of retention used by each institution can depend upon their unique profile and mission. The results of this research may have implications for colleges offering a program mix that includes certificate, diploma, applied degree, and university transfer options. The retention rates produced by this research may be used as internal benchmarks by the two institutions who

supplied the data; they may therefore be used to demonstrate quality and measure future progress toward goals.

The results may have implications for quality assessment and improvement at the colleges that supplied data for the current study. Estimates of retention rates may be used for external comparison purposes; in this case the results for both institutions were comparable to published estimates. Estimates of retention may also be used to develop internal benchmarks across years. Comparison of annual estimates may be used to measure continual improvement of quality, and to assess the effects of SEM strategies on retention. Examination of retention rates by credential may also identify sub-groups of the student population who are less likely to be retained. The results of this research could be used to inform SEM strategies that may increase retention among at-risk sub-groups.

The results of this research may also have implications for the institutions involved to use the data to demonstrate return on investment of public funds and alignment with public policy objectives (Bogue & Dandridge Johnson, 2010). With current accountability frameworks and performance based funding becoming more openly discussed in several provinces such as Alberta, Ontario, and Quebec (Alberta Advanced Education, 2005; Alberta Enterprise and Advanced Education, 2013; Council of Ontario Universities, 2010; Pakravan, 2006; Shanahan & Jones, 2007), institutions are being forced to demonstrate both quality and return on investment of public funds with indicators such as retention rates.

Society

Society as a whole benefits indirectly from greater post-secondary attainment among citizenry (Canadian Council on Learning, 2009a). Post-secondary attainment among citizenry generally produces benefits to society as a whole, and aligns with both academic-humanist and

economic-utilitarian societal values (Kirby, 2007). Economic and demographic changes have increased the need for participation in post-secondary education in order to meet labour market demand and ensure the nation's prosperity in a knowledge based economy (Alberta Enterprise and Advanced Education, 2013; Kirby, 2007; Shanahan & Jones, 2007). Globalization has necessitated a more utilitarian, market-oriented approach to PSE that focusses on developing and supporting a knowledge-based economy (Kirby, 2007; Shanahan & Jones, 2007).

Society, represented by governing bodies, typically strives to ensure that that the post-secondary system is efficiently meeting policy objectives and public expectations (Alberta Advanced Education and Technology, 2007; 2006a; 2005; Maritime Provinces Higher Education Commission, 2013; Ministry of Advanced Education, British Columbia, 2013; Ministry of Advanced Education, Saskatchewan, 2012). Alignment with public policy objectives can be measured by the outcomes of graduates (e.g. graduate employment and skill levels). From a societal perspective retention is an important stepping stone to graduation and employment. Retention is inherently measured through graduation rates, since graduation cannot occur unless a student is retained.

The results of the current study are potentially valuable to society because they provide a unique contribution to our understanding of the dynamics of retention among an under-studied group - Canadian community college students. Although the implications of this research for society are indirect, the results are potentially valuable if used to increase the likelihood that students successfully participate in higher education beyond the critical first semester of studies. The recommendations produced from this research suggest multiple ways in which student retention can be increased beyond this critical time period. Every consecutive semester of

enrolment can be considered an additional step towards a more educated citizenry, which in turn will help fulfill labour market demand and support a knowledge-based economy.

Recommendations

One of the primary aims of this research was to provide information on the dynamics of retention among Canadian college students that could be used by stakeholders. Due to the paucity of available research, this study attempted to add to our understanding of retention dynamics among college students within the Canadian context. The results of this study may help guide policy and practice among stakeholders. The results may also add to the foundation of existing research to help inform the direction of future research. The following section will present recommendations for stakeholders and for future research.

Recommendations for stakeholders

There have been relatively few studies that have examined the dynamics of retention among Canadian college students. Therefore, the results of this research add to the small body of published literature on the topic. There are three main implications of this research for stakeholders. First, the results of this research indicate that, in general, the predictors of retention among this sample are fairly similar across institutions and credential types. Second, despite this consistency, institutional differences in the magnitude of the effects of these predictors should be considered. Third, the dynamics of retention may differ for certain sub-groups of students. Each of these implications will be discussed below.

The recommendations provided here should be headed with caution. The findings and related recommendations produced from this research should be used to *inform* but not *drive* decision making at individual institutions. The recommendations must be considered within the institutional context to inform policy and practice, and the appropriateness and feasibility of

related strategies should be evaluated collectively by stakeholders. The limitations of this research should be considered in conversations among stakeholders as part of this evaluative process.

Consistency of results

Overall, enrolment in full-time studies and having a GPA above the 50th percentile were the strongest predictors of retention in this analysis. In all six of the multivariable models, full-time enrolment (defined as enrolment in at least nine credits) predicted retention. Having a GPA above the 50th percentile predicted retention in five of the six models. The consistency of these results suggests that they should be considered key factors of interest to stakeholders concerned with increasing retention among college students.

From an institutional perspective, implementing policies and practices (possibly through SEM processes) aimed at encouraging full-time enrolment, and supporting students' academic performance, may increase retention rates in the future. Staff who typically have direct contact with prospective students (e.g. recruiters, academic advisors, faculty) should be made aware of the potential impact that full-time enrolment may have on retention so that credit load can be discussed with the students. Once in contact with staff, students might be encouraged to seek academic advising, or conversely intrusive advising (outreach) could be employed to ensure that, where possible, students enroll in at least nine credits in the first semester. Appropriate supports may be appropriate for some students with greater life demands in order to facilitate full-time enrolment.

Institutional SEM strategies may also be informed by the results regarding the effect of GPA on retention. Administrators and staff may find value in reviewing SEM strategies that support academic achievement. Administrators and staff are encouraged to conduct an audit of

the academic supports currently available to students within their colleges, including consideration of student awareness/utilization of available supports. The audit may also include student tracking in order to evaluate the success of each support service. Student grades before and after accessing supports, along with surveys for student feedback on support effectiveness may be helpful in benchmarking and improving academic resources available to students. Expansion or promotion of supports that contribute to improved academic performance among students may lead to an improved within-year retention rate.

Prospective students may benefit the most by focusing on environmental factors associated with retention over which they can exert some control. More specifically, these environmental factors include credit load and GPA. Prospective students may want to consider enrolling in full-time as opposed to part-time studies when possible in order to increase their odds of retention. Seeking academic advising prior to enrolment may help students determine the appropriate credit load to suit their unique needs and lifestyles. Seeking advising may also make prospective students aware of supports and services that could make full-time enrolment possible while juggling competing life demands. While balancing life demands is important, enrolling in at least nine credits in the fall semester may improve the odds of retention in the first semester for some students.

Students should also be aware of and apply strategies to improve academic performance in the first semester of studies. Such strategies may include participation in orientation programs, study skills workshops, peer tutoring, attending all classes, completion of course readings, and attending instructors' office hours, among others. Having an understanding of such strategies, and an awareness of the institutional resources available to support students, may be helpful.

Institutional differences in the magnitude of effects

Despite the overall consistency of the independent variables identified as predictors of retention in this research, there was substantial variation in the magnitude and precision of the estimates within and between the two institution-specific models. For staff working at these institutions, the results indicate that the development of targeted strategies to increase retention may require an individualized understanding of the data and context-specific approaches. Understanding the magnitude of the effect of covariates within each institution-specific model may suggest areas where greater emphasis should be placed when developing strategies to increase retention.

College administrators and staff may find it helpful to consider the magnitude of effects of factors that predict retention within their institution when developing targeted programs to increase retention. Understanding the magnitude of one effect compared to another may provide guidance as to where resource allocations may produce the greatest return on investment. In times of resource scarcity and outcomes based funding, allocation of resources towards strategies that may produce better returns (in terms of increased retention rates) may be preferable.

Retention dynamics among sub-groups

Despite the overall consistency in the aggregate results, the independent variables shown to predict retention differed by credential. Stakeholders may find value in understanding the unique factors that may influence retention among sub-groups of students enrolled in different credentials. The results pertaining to each multivariable model in the results of sub-question 1.e presented specific factors that predicted retention among students enrolled in university transfer, parchment, and upgrading/non-parchment programs.

Several differences existed between the predictors of retention among the three credential types. Full-time credit load consistently predicted retention among students enrolled in university transfer programs, parchment programs, and upgrading/non-parchment programs. Administrators and staff at these colleges are encouraged to observe the importance of the effect of credit load on retention in the first semester. Again, staff at colleges should advise students on the potential benefit of full-time enrolment, and encourage awareness and utilization of academic support services. Prospective students may find value in enrolling in full-time as opposed to part-time studies if possible, regardless of credential-type.

Having a GPA above the 50th percentile predicted retention among students enrolled in university transfer programs and parchment programs, but not among students enrolled in upgrading/non-parchment programs. Administrators and staff at institutions are advised to consider developing and implementing strategies to enhance academic supports specifically among students in university transfer and parchment programs. As mentioned above, this may include an audit of academic supports, focusing on awareness, utilization, and effectiveness. Prospective students considering enrolment in university transfer or parchment programs are encouraged to become aware of the academic supports available prior to enrolment. In addition, understanding when and how to access these services may be helpful in increasing the odds of retention within the first semester.

Retention among students enrolled in university transfer programs may also be influenced by the number of hours spent working per week. The results of the current study show that student who did not work, and those who worked less than or equal to ten hours per week were more likely to be retained compared to those who worked more than 20 hours per week. College administrators and staff are encouraged to identify the reasons for employment among

students enrolled university transfer programs (e.g. financial need, experience, etc.), and to implement strategies that might encourage these students to work at a job as little as possible while attending college. For example, if the majority of university transfer students are discovered to be working out of financial necessity, increased availability of scholarships specifically targeted at university transfer students in financial need may allow some students to reduce the number of hours spent working per week.

Conversely, students enrolled in university transfer programs are encouraged to consider working as little as possible while attending college, if possible. Taking measures to reduce financial stress (possibly through student loans or scholarships) may allow them to focus more fully on their studies.

The current study provides new information about predictors of retention among students enrolled in upgrading and other non-parchment programs. Retention among students enrolled in upgrading/non-parchment programs was predicted by the expressed commitment to the goal of graduation. College staff who have contact with students are advised to encourage these students to access academic advising. Academic advisors may help students set and measure progress towards academic goals. Academic advisors may be able to offer suggestions that may help students meet their goals faster. Similarly, students considering enrolment in upgrading or other non-parchment program are advised to seek academic advising to help with academic goal setting upon enrolment.

In addition to considering the predictors of retention by credential, it may be important to consider the magnitude of effects within each credential type when developing targeted programs to increase retention. Staff at colleges who are interested in improving the odds of retention among students may find value in understanding differences in the relative effects of each factor

before developing such programs. Understanding the magnitude of one effect compared to another may provide guidance as to where resource allocations may produce the greatest return on investment.

Recommendations for future research

This study sought to examine the dynamics of retention among Canadian college students. In doing so, within-year retention was examined among a single cohort of students attending the first year of a program at two colleges in Western Canada. Recommendations for future research generally fall into three categories: those related to the sample used, those relating to the independent variables examined, and those relating to the design.

Sample

Future research examining retention among all students - including those enrolled in years 2-4 of study – would be of value. As suggested by Jorgensen et al. (2009), the rate and dynamics of retention may change over students' academic careers. Further differences over time among sub-groups of students may also be realized.

In this research, several exclusion criteria were applied. Students enrolled in ESL and apprenticeship programs were not included. These students were excluded due to the timing of intake and their unique reasons for attending. Examination of retention among these students in future research may be helpful; and may potentially identify similarities and differences in the dynamics of retention that could inform policy and practice.

In the current study, students who withdrew from their studies prior to the add/drop deadline were not included because they would have incomplete data on environmental variables. This would be particularly true of GPA. However, these students represent true departures from an institution. Inclusion of students who withdraw prior to the add/drop deadline

may produce valuable information in future research. It is recommended that these students be included in future research, and that GPA be imputed to equal zero.

Although the analyses conducted in the current study had adequate power to detect true effects, the precision of the estimates in several of the models suggested that a larger sample size would be helpful. Gathering multiple years of data, or from more than two colleges, may ensure that the estimates of the effects are more precise.

Independent variables

This research included a variety of student input and environmental factors collected from a combination of student records and survey items. In the case of certain variables – namely measures of past academic performance – future research should make every effort to collect from student records. In particular, high school average and whether or not a student passed grade 12 pure math may be of particular importance in order to avoid the potential for bias.

Potential bias aside, the results of the current study suggest that first semester GPA may be a more important predictor of retention than past academic performance. Future research examining retention among Canadian college students should include first semester GPA as an independent variable.

Future research should also further examine the influence of age and gender on retention. The results of the current study do not support those in published literature, and further research is needed to clarify how these demographic factors influence retention. At the very least, age and gender should be considered as potential modifying factors in multivariable models of retention. Even if direct associations with retention are not observed, age and gender are likely associated with a number of other independent variables.

Data were not available for this study to evaluate the effect of student engagement on retention. Engagement/integration has been noted to impact retention among Canadian college students (Fisher and Engemann, 2009; Kirby and Sharpe, 2001, Jorgensen, 2009) and should be included in future research if possible. Similar to Sarkar (1993), the results of this study indicate that goal commitment may influence retention among certain students, but this also needs to be further explored by future research.

Design

This research used multivariable logistic regression to examine the dynamics of retention. No other Canadian studies of college student retention that used multivariable logistic regression were identified in the literature, therefore the results could not adequately be confirmed or refuted. The results of the current study need to be confirmed/refuted by future research using a similar design.

The design used in the current study was also exploratory in nature. This limits the generalizability of the data to dissimilar situation. Future research performing confirmatory regression analysis may provide further direction on how the results could be used outside of the institutions that supplied data.

The current study made no attempt to assess for confounding or effect modification, even though associations between covariates were considered in interpretation of the results. Future research using a similar design should examine the role of confounding and effect modification in order to improve our understanding of the dynamics of retention.

This research identified that the dynamics of retention may differ by sub-groups of students (by credential). These results need to be confirmed by future research. It may also be interesting to examine the dynamics of retention among more specific sub-groups of students.

Within the parchment programs, differences may be delineated if examined by certificate, diploma, and applied degree programs. Similarly, other sub-groups not identified in the current study may also exist, for whom the dynamics of retention are still unknown.

Conclusions

The overarching goal of this research is to produce information that institutions can use as a basis for improving the success and retention of the students that they admit, register and convocate. This study sought to estimate the rate of retention and to examine the dynamics of retention among a sample of Canadian college students. The results of this research suggest that the estimated rate of retention among Canadian college students is congruent with previous published estimates. The results also suggest that two environmental factors – credit load and GPA – are the main determinants of retention. However, certain sub-groups of students exist for whom the dynamics of retention are unique.

While this research has several strengths and weaknesses, it adds to the sparse body of literature examining the dynamics of retention among Canadian community college students. The results of the research have implications for those who have a vested interest in post-secondary education in Canada, including students, college administrators and staff, and society in general. The recommendations put forth from this research may inform policy and practice at colleges, and may help students make informed choices about their education. These recommendations may help improve retention among future students.

Although in the United States retention is already a well-studied and observed quality indicator among both institutions and the governments (Bogue & Dandridge Johnson, 2010; Ross, Kena, Rathbun, Kewal Ramani, Zhang, Kristapovich, et al., 2012; U.S. Department of Education, 2011), it is relatively unexplored in the Canadian context. This is changing, however,

and in the future we may see greater interest in persistence among both Canadian institutions and provincial ministries. As the Canadian PSE system becomes more complex, and as external forces such as demographics and the economy shift, strategies to increase student persistence may become an increasingly important to maintain enrolment and ensure sufficient graduates are being produced to support the economy. Enhancing our understanding of retention in the Canadian college context, and taking action to improve retention, may contribute to Canada's future prosperity in a knowledge based economy.

References

Alberta Advanced Education (2005). *A Learning Alberta. Advanced Education: A Cross Jurisdictional Overview of Accessibility, Affordability, and Quality*. Edmonton, AB: Author. Retrieved from http://eae.alberta.ca/media/134906/cjoverview.pdf

Alberta Advanced Education (2006). *A Learning Alberta. Recommendations from Transforming the Advanced Education System Subcommittee*. Retrieved from http://eae.alberta.ca/media/133979/transforming_the_advanced_education_system.pdf

Alberta Enterprise and Advanced Education (2012). *Campus Alberta Planning Resource.* Retrieved from http://eae/alberta.ca/capr

Alberta Enterprise and Advanced Education (2013, September). *Introduction to Results-Based Budgeting.* Presented at the meeting of Alberta Institutional Researchers, Calgary, Alberta.

Alberta Enterprise and Advanced Education (2013). Campus Alberta Planning Resource. Retrieved from http://eae.alberta.ca/capr.

Alberta Advanced Education and Technology (2007). *Roles and Mandates Policy Framework for Alberta's Publically Funded Advanced Education System.* Retrieved from http://eae.alberta.ca/media/133783/rmpf.pdf

American Association of Community Colleges. (2000). *National profile of community colleges: Trends and statistics, 1997-1998*. Washington, DC: American Association of Community Colleges.

Askari, M., Lee, W., Chevannes, V., Marlin, S., & Blair, M. (2011, July 21). Re: The effect of funding cuts on higher education [Online forum comment]. Retrieved from https://blackboard.ucalgary.ca/webapps/portal/frameset.jsp?tab_id=_2_1&url=%2fwebap

ps%2fblackboard%2fexecute%2flauncher%3ftype%3dCourse%26id%3d_91479_1%26url%3d

Aslanian, C. B. (2001). *Adult students today*. New York: The College Board.

Association of Canadian Community Colleges (2008). Pan-Canadian Study of First-Year College Students. Retrieved from http://www.accc.ca/english/publications/reports/student_study.htm

Astin, A.W. (1975). *Preventing students from dropping out.* San Francisco: Jossey-Bass.

Astin, A.W. (1984). Student involvement: A developmental theory for higher education. *Journal of College Student Personnel*, 25, 297-308.

Astin, A.W. (1991). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education*. New York: Macmillan.

Astin, A.W. (1993). *What matters in college?* San Francisco: Jossey-Bass.

Astin, A.W. & Antonio, A.L. (2012). *Assessment for Excellence* (2nd ed.). New York: Rowman & Littlefield Publishers, Inc.

Bean, J.P. (1990) Why students leave: Insights from research. In D. Hossler, & J.P. Bean (Eds.), *The strategic management of college enrollments*. San Francisco: Jossey-Bass.

Bean, J.P., & Metzner, B.S. (1985) A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research*, Vol. 55, No. 4, p. 485.

Bennet, E. (2010). Validation experiences and persistence among community college students. *The Review of Higher Education*, 34(2), 193-230.

Berger, J. B. (2000). Optimizing capital, social reproduction, and undergraduate persistence. In J. M. Braxton (Ed.), *Reworking the student departure puzzle* (pp. 95-124). Nashville, TN: Vanderbilt University Press.

Bers, T. (1988). Student major choices and community college persistence. Research in Higher Education, 29(2), 161-173.

Black, J. (2010). Strategic enrolment issues facing higher education leaders. In J. Black (Ed.), *Strategic Enrolment Intelligence* (pp. 15-26). Retrieved from http://www.academica.ca/strategic-enrolment-intelligence

Blair, M. (2009). *Prevalence and Correlates of Neurocognition in Older Patients Presenting for Coronary Care* (Master's thesis). University of Calgary, Calgary, Alberta, Canada.

Bogue, E.G. (1998). Quality assurance in higher education: The evolution of systems and design ideals. *New Directions for Institutional Research*, 99, 7-18.

Bogue, E.G., & Dandridge Johnson, B. (2010). Performance incentives and public college accountability in the United States: A quarter century policy audit. *Higher Education Management and Policy*, 22(2), 9-30.

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241-258). New York: Greenwood.

Bourdieu, P., & Passeron, J. C. (1977). *Reproduction in education, society, and culture*. Beverly Hills, CA: Sage.

Braxton, J.M., Vesper, N. & Hossler, D. (1995). Expectations for college and student persistence. *Research in Higher Education*, 36(5), 595-612.

Braunstein, A., McGrath, M., & Pescatrice, D. (2000/2001). Measuring the impact of financial factors on college persistence. *Journal of College Student Retention*, 2(3), 191-203.

Cabrera, A. F. (1994). Logistic regression analysis in higher education: An applied perspective. In J. C. Smart (Ed.), Higher Education: Handbook of Theory and Research (pp. 225-56). New York: Agathon Press.

Canadian Council on Learning (2005). Lessons in learning, October 28, 2005. Vancouver, BC: Canadian Council on Learning.

Canadian Council on Learning (2006). *Report on Learning in Canada 2006, Canadian Post-Secondary Education: A Positive record – An Uncertain Future*. Ottawa, Canada. Retrieved from www.ccl-cca.ca

Canadian Council on Learning (2009a). Up to par: *The Challenge of Demonstrating Quality in Canadian Post Secondary Education. Challenges in Canadian Post Secondary Education*. Ottawa, Canada. Retrieved from www.ccl-cca.ca

Canadian Council on Learning (2009b). *Post-Secondary Education in Canada: Meeting our Needs?* Ottawa, Canada. Retrieved from www.ccl-cca.ca

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, and Social Sciences and Humanities Research Council of Canada, Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, December 2010. Retrieved from http://www.pre.ethics.gc.ca/pdf/eng/tcps2/TCPS_2_FINAL_Web.pdf

Chagnon, J., & Milan, A. (2011). Population growth: Canada, provinces, and territories, 2010. Retrieved from http://www.statcan.gc.ca/pub/91-209-x/2011001/article/11508-eng.htm

Cofer, J. (1998). *A decade of indecision: The effects of federal student financial aid policy from 1987 to 1996 on within-year persistence of four-year undergraduate students*. Unpublished doctoral dissertation, University of Arkansas at Little Rock, Little Rock, AR.

Cofer, J., & Somers, P. A. (1997, October). *Mortgaging their future: Debtload and undergraduate persistence*. Paper presented at the Association for the Study of Higher Education Annual Meeting, Albuquerque, NM.

Cofer, J., & Somers, P. A. (1998, April). *Debtload and undergraduate persistence: A comparison of public and private college students*. Paper presented at the American Educational Research Association Meeting, San Diego, CA.

Cofer, J., & Somers, P. (2000).Within-year persistence of students at two-year colleges. Community College Journal of Research and Practice, 24, 785-807.

Cohen, A. M., & Brawer, F. B. (1996). *The American community college* (3rd ed.). San Francisco: Jossey-Bass.

Coleman, J.S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, 94, S95-S120.

Council of Ministers of Education, Canada (1999). *A Report on Public Expectations of Postsecondary Education in Canada.* Retrieved from http://www.cmec.ca/Publications/Lists/Publications/Attachments/22/expectations.en.pdf

Council of Ontario Universities (2010). *Framework for Planning and Funding of Enrolment.* Retrieved from http://www.cou.on.ca/publications/reports/pdfs/june-2010---framework-for-planning-and-funding-of-

Crawford Sorey, K., & Harris Duggan, M. (2008). Differential predictors of persistence between community college adult and traditional-aged students. *Community College Journal of Research and Practice*, 32(2), 75-100.

Dale, M. (2010. *Trends in Age Composition of College and University Students and Graduates.* Statistics Canada. Retrieved from http://www.statscan.gc.ca/pub/81-004-x/2010005/article/11386-eng.htm

Diallo, B., Trottier, C. & Doray, P. (2009). What do we know about pathways and transitions of

Canadian students in post-secondary education? Note 1: Transitions Project, Number 46. The Canadian Millennium Scholarship Foundation. Retrieved from www.millenniumscholarships.ca

Dooley, M.D., Payne, A.A., & Robb, A.L. (2011). *Understanding the Determinants of Persistence and Academic Success in University: An Exploration of Data from Four Ontario Universities.* Higher Education Quality Council of Ontario. Retrieved from http://www.heqco.ca/en-CA/Research/Research%20Publications/Pages/Home.aspx

Dougherty, K. (2003). The community college: The origins, impacts, and futures of a contradictory institution. In J. Ballantine & J. Spade (Eds.), *Schools and Society* (pp. 377-385). Belmont, CA: Wadsworth.

Council on Learning.

Fike, D.S., and Fike, R. (2008). Predictors of first year retention in the community college. *College Review*, 36(2), 68-88.

Finnie, R. & Qiu, H. (2008). *The Patterns of Persistence in Post-Secondary Education in Canada: Evidence from the YITS-B dataset.* Retrieved from www.mesa-project.org

Fisher, R. & Engemann, J. (2009). Factors affecting attrition at a Canadian College. Canadian Council on Learning. Vancouver, BC: Canadian Council on Learning.

Flynn, W.J.; & Vredevoogd, J. (2010). The future of learning. 12 views on emerging trends in higher education. *Planning for Higher Education*. 38(2), 4-10.

Grayson, P.J. (1998). Racial origin and student retention in a Canadian University. *Higher Education*, 36, 323-252.

Grayson, J.P., & Grayson, K. (2003). Research on retention and attrition. Montreal: Canadian Millennium Scholarship Foundation.

Green, T. (2013). Retention plans: Why they fail and how we can fix them. Strategic Enrollment Management Quarterly, 1(2), 113-122.

Hagedorn, L.S. (2005). How to define retention: A new look at an old problem. In A. Seidman (Ed.), *College student retention* (pp.89-105). Westport: Praeger Publishers.

Hand, C., and Payne, E.M. (2008). First-generation college students: A study of Appalachian student success. *Journal of Developmental Education*, 32(1), 4-15.

Hand, D., Mannila, H., & Smyth, P. (2001). *Principles of Data Mining*. Cambridge, MA: MIT Press.

Hango, D. (2011). *Delaying Post-secondary Education: Who delays and for How Long?* Statistics Canada. Retrieved from http://www.statscan.gc.ca/pub/81-595-m/81-595-m2011090-eng.pdf

Henderson, S. (2012). Integrating evolving perspectives: The roots and wings of strategic enrollment management. In B. Bontrager, D. Ingersoll, & R. Ingersoll (Eds.), *Strategic Enrollment Management: Transforming Higher Education.* Washington, DC: American Association of Collegiate Registrars and Admissions Officers.

Herzog, S. (2003). Measuring determinants of student return vs. Transfer vs. stopout vs. dropout: A first-to-second year analysis of new freshmen. Paper presented at the meeting of the California Association for Institutional Research, Rohnet Park, California.

Herzog, S. (2005). Measuring determinants of student return vs. dropout/stopout vs. transfer: A first-to-second year analysis of new freshmen. *Research in Higher Education,* 46(8), 883-928.

Higher Education Quality Council of Ontario. (2013). *Quality: Shifting the Focus. A Report from the Expert Panel to Assess the Strategic Mandate Agreement Submissions*. Toronto: Higher Education Quality Council of Ontario.

Hirji, K. K. (2001). Exploring data mining implementation. *Communications of the ACM,*

44(7), 87-93.

Hippensteel, D. G., St. John, E. P., & Starkey, J. B. (1996). Influence of tuition and student aid on within-year persistence by adults in 2-year colleges. *Community College Journal of Research and Practice*, *20*, 233–242.

Hosmer, D.W., and Lemeshow, S. (2000). Applied Logistic Regression (2nd Ed). New York, NY: John Wiley & Sons, Inc.

Hossler, D., and Kalsbeek, D. (2013). Enrollment management and managing enrollments: Revisiting the context for institutional strategy. *Strategic Enrollment Management Quarterly,* 1(1), 5-25.

Howe, N., & Strauss, W. (2000). *Millennials Rising: The Next Generation*. New York: Vintage Books.

Ishitani, T.T. (2003). A longitudinal approach to assessing attrition behavior among first-generation students: Time-varying effects of pre-college characteristics. *Research in Higher Education,* 44(4), 433-449.

Jorgensen, S., Ferraro, V., Fichten, C., & Havel, A. (2009). *Predicting College Retention and Dropout: Sex and Disability*. Adaptech Research Network. Retrieved from http://www.adaptech.org/pubs/abPredictingCollegeRetentionAndDropoutE.pdf

Joseph, M., Yakhou, M., & Stone, G. (2005). An educational institution's quest for service quality: Customers' perspective. *Quality Assurance in Education*, 13(1), 66-82.

Kirby, D. (2007). Reviewing Canadian post-secondary education: Post-secondary education policy in post-industrial Canada. *Canadian Journal of Educational Administration and Policy*, 65, 1-24.

Kirby, D. and Sharpe, D. (2001). Student attrition from Newfoundland and Labrador's public college. *Alberta Journal of Educational Research*, 47(4), 353-368.

Larose, D. T. (2005). *Discovering Knowledge in Data*. Hoboken, NJ: John Wiley & Sons, Inc.

Liu, R. (2010). Alienation and first-year student retention. *Association for Institutional Research Professional File,* 116, 1-18.

Ma, X. & Frempong, G. (2008). *Reasons for non-completion of postsecondary education and profile of postsecondary dropouts.* Ottawa, ON: Learning Policy Directorate, Human Resources and Social Development Canada.

Madgett, P.J., and Belanger, C.H. (2008). First university experience and student retention factors. *Canadian Journal of Higher Education,* 38(3), 7-96.

Maritime Provinces Higher Education Commission. (2013). *Students at the Heart: Stakeholder Reactions to the MPHEC's Proposed Approach to Quality Assurance at Maritime Universities*. Retrieved from http://www.mphec.ca/media/48926/QA_Consultation_Report_EN.pdf

Marshall, C. (2008). The school of language and liberal arts writing program and retention at Fanshawe College. *College Quarterly*,11 (2). Retrieved from http://www.collegequarterly.ca/2008-vol11-num02-spring/marshall.html

McCabe, R. H. (2000). *No one to waste: A report to public decision-makers and community college leaders*. Washington, DC: Community College Press.

McCabe, R. H. (2000). *No one to waste: A report to public decision-makers and community college leaders*. Washington, DC: Community College Press.

McClintock, P.J.; & Snider, K.J.G. (2008). Driving decision making with data. *New Directions for Institutional Research,* 137, 15-40.

McDonough, P. M. (1997). *Choosing colleges: How social class and schools structure*

opportunity. Albany: State University of New York Press.

McElroy, L. (2005). Student aid and university persistence: Does debt matter? Montreal: Canada Millennium Scholarship Foundation.

Menard, S. (1995). Applied logistic regression. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-106. Thousand Oaks, CA: Sage.

Ministry of Advanced Education, British Columbia. (2013). *Annual Service Plan*. Retrieved from http://www.bcbudget.gov.bc.ca/2013_June_Update/sp/pdf/ministry/aved.pdf

Ministry of Advanced Education, Saskatchewan. (2012). *Plan for 2013-14*. Retrieved from http://www.ae.gov.sk.ca/ae-plan-2013-14

Ministry of Education and Training, Ontario. (1996). *Future Goals for Ontario Colleges and Universities.* Retrieved from http://www.tcu.gov.on.ca/eng/document/discussi/postdeng.html#1.1

Ministry of Training, Colleges and Universities. (2005). Ontario: A Leader in Learning. Report and Recommendations. Retrieved from http://www.tcu.gov.on.ca/eng/document/reports/postsec.pdf

Nguyen, A., Hays, B., & Wetstein, M. (2010). Showing incoming students the campus ropes: Predicting student persistence using a logistic regression model. *Journal of Applied Research in the Community College*, 18(1), 16-21.

Nicoletti, A. & Merriman, W. (2007). Teaching Millenial. *Momentum*, April/May, 28-31.

Norman, G.F., & Streiner, D.L. (2000). *Biostatistics: The Bare Essentials*. Hamilton: B.C. Decker.

Pakravan, P. (2006). The Future is Not What it Used to Be. Re-examining Provincial Postsecondary Funding Mechanisms in Canada. *C.D. Howe Institute*. Retrieved from www.cdhowe.org

Parkin, A. & Baldwin, N. (2009). Persistence in post-secondary education in Canada: The latest research. Note 8: Transitions Project. The Canadian Millennium Scholarship Foundation. Retrieved from www.millenniumscholarships.ca

Patterson, P., Vanbalkom, D., Jensen, D., Cummings, H. (2009). *Higher Education Leadership Issues – 2009. Survey of Current issues in Canadian higher education. Perspectives from the leaders.* Unpublished manuscript. University of Calgary, Calgary, Canada.

Petrie, A., & Sabin, C. (2005). *Medical Statistics at a Glance*. Malden, Massachusetts: Blackwell Publishing.

Pittman, K. (2008). Comparison of Data Mining Techniques used to Predict Student Retention (Doctoral Dissertation). Retrieved from ProQuest.

Rae, B. (2005). *Ontario: A Leader in Learning*. Retrieved from http://www.tcu.gov.on.ca/eng/document/reports/postsec.pdf

Reason, R.D. (2009). Student variables that predict retention: Recent research and new developments. *NASPA Journal*, 46(3), 482-501.

Roberts, J. & Styron, R. (2010). Student satisfaction and persistence: Factors vital to student retention. *Research in Higher Education*, 6, 1-18.

Robertson, L.J., & Taylor, C.A. (2009). Student persistence in the human sciences: Freshman to sophomore year. *Journal of Family and Consumer Services*, 101 (1), 36-43.

Ross, T., Kena, G., Rathbun, A., Kewal Ramani, A., Zhang, J., Kritapovich, P. & Manning, E. (2012). *Higher Education: Gaps in Access and Persistence Study.* US Department of Education, National Center for Education Statistics. Washington, DC: Government Printing Office. Retrieved from http://nces.ed.gov/pubs2012/2012046.pdf

Rothman, K.J. (2002). *Epidemiology: An Introduction*. New York: Oxford University Press.

Sarkar, G. (1993). SIAST Retention Study. Factors affecting retention of first-year students in a Canadian technical institute of applied science and technology. Paper presented at the Canadian Institutional Researchers and Planners Conference, Vancouver, British Columbia.

Settle, J.S. (2011). Variables that encourage students to persist in community colleges. Community College Journal of Research and Practice, 35(4), 281-300.

Shanahan, T. & Jones, G.A. (2007). Shifting roles and approaches: Government coordination of post-secondary education in Canada, 1995-2006. *Higher Education Research & Development*, 26(1), 31-43.

Skinkle, R. (2010). Actionable intelligence: Research for SEM. In J. Black (Ed.), *Strategic Enrolment Intelligence* (pp. 51-68). Retrieved from http://www.academica.ca/strategic-enrolment-intelligence

Smith, C., & Gottheil, S. (2008). Enrollment or enrolment. *College and University*, *84*(2), 28-38. Retrieved from OmniFile Full Text Select database.

St. John, E. P., & Starkey, J. B. (1994). The influence of costs on persistence by traditional college-age students in community colleges. *Community College Journal of Research and Practice, 18*, 201–213.

Statistics Canada (2010). The financial impact of student loans. *Perspectives on Labour and Income,* 11(1), 1-18.

Statistics Canada. (2010). *Population Projections for Canada, Provinces, and Territories. 2009-2036.* Retrieved from http://www.statcan.gc.ca/pub/91-520-x/91-520-x2010001-eng.pdf

Steele, K. (2010). The changing Canadian PSE landscape. In J. Black (Ed.), *Strategic Enrolment*

Intelligence (pp. 27-50). Retrieved from http://www.academica.ca/strategic-enrolment-intelligence

Swail, W.S., Redd, K.E., & Perna, L.W. (2003). Retaining minority students in higher education. *ASHE-Eric Higher Education Report*, 30(2), 1-187.

Tagg, J. (2008). Changing minds in higher education: Students change, so why can't college? *Planning for Higher Education*, 37(1), 15-22.

Tinto, V.(1993). Leaving College. Chicago: University of Chicago Press.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research* 45, 89-125.

United Nations (1948). *Universal Declaration of Human Rights.* New York: Author. Retrieved from http://www.un.org/en/documents/udhr/index.shtml

U.S. Department of Education (2011). *Six-Year Attainment, Persistence, Transfer, Retention, and Withdrawal Rates of Students Who Began Postsecondary Education in 2003-04.* Retrieved from http://nces.ed.gov/pubs2011/2011152.pdf

Vorhees, R. A. (1987). Toward building models of community college persistence: A logistic analysis. *Review of Higher Education*, 26 (2), 115–29.

Wells, R. (2008). The effects of social and cultural capital on student persistence: Are community colleges more meritocratic? Community College Review, 36, 25-46.

Yale, A. (2010). Expanding the Conversation about SEM: Advancing SEM Efforts to Improve Student Learning and Persistence—Part II. *College and University*, *85*(4), 28-38. Retrieved from OmniFile Full Text Select database.

Appendix A: Factors Associated with Retention among College and University Samples

Input Factors	Source
Age	Settle, 2011; Fike & Fike, 2008; Cofer & Somers, 2010; Nguyen et al., 2010; Bers, 1988; Robertson & Taylor, 2009; Reason, 2009; Sarkar, 1993
Gender	Grayson, 1998; Bers, 1988; Reason, 2009; Sarkar, 1993; Ma & Frempong, 2008
Direct entry from high school	Settle, 2011
Parental education/ first generation student	Fike & Fike, 2008; Nguyen et al., 2010; Wells, 2010; Hand & Payne, 2008; Grayson, 1998; Ishitani, 2003
Race/ethnicity	Nguyen et al., 2010; Liu, 2010; Reason, 2009; Grayson, 1998; Sarkar, 1993
Place of origin	Herzog, 2003; Herzog, 2005
Marital Status	Reason, 2009; Sarkar, 1993; Ma & Frempong, 2008
Dependent children	Sarkar, 1993
Academic aspirations	Cofer & Somers, 2010; Sarkar, 1993; Ma & Frempong, 2008
High school performance/diploma	Cofer & Somers, 2010; Sarkar, 1993; Kirby & Sharpe, 2001; Ma & Frempong, 2008
Academic confidence	Robertson & Taylor, 2009
Academic goal commitment	Sarkar, 1993
Drug abuse in high school	Ma & Frempong, 2008
Passing developmental course(s) / test scores / academic preparedness	Fike & Fike, 2008; Nguyen et al., 2010; Herzog, 2003; Fisher and Engemann, 2009; Marshall, 2008; Kirby & Sharpe, 2001
Environmental Factors	*Source*
Contact with faculty	Settle, 2011; Roberts & Styron, 2010; Bennet, 2010
GPA	Settle, 2011; Grayson, 1998; Cofer & Somers, 2010; Nguyen et al., 2010; Liu, 2010; Braunstein, McGrath, & Pescatrice, 2000/2001; Reason, 2009; Herzog, 2005; Ma & Frempong, 2008
Dropping courses in first semester	Fike & Fike, 2008;

Input Factors	Source
Credit load in first semester/Ft vs. PT	Fike & Fike, 2008; Cofer & Somers, 2010; Settle, 2011; Herzog, 2005; Kirby & Sharpe, 2001
Enrolment in orientation course	Nguyen et al., 2010; Derby & Smith, 2004
Number of counseling appointments in 1st year	Nguyen et al., 2010;
Major	Bers, 1988
Receptivity to academic assistance	Robertson & Taylor, 2009
Study hours while in school	Madgett & Belanger, 2008
Friends attending same institution/social connectedness	Settle, 2011; Roberts & Styron, 2010; Wells, 2008;
Family emotional support	Robertson & Taylor, 2009
Living arrangements	Settle, 2011; Madgett & Belanger, 2008; Liu, 2010; Herzog, 2005; Grayson, 1998;
Other family members attending college	Settle, 2011
Student engagement/ social & academic integration	Fisher and Engemann, 2009; Kirby & Sharpe, 2001; Ma & Frempong, 2008
Working while in school	Grayson, 1998; Sarkar, 1993
Family resources/Socioeconomic status	Wells, 2008; Cofer & Somers, 1998; Braunstein, McGrath, & Pescatrice, 2000/2001; Herzog, 2003; Ma & Frempong, 2008
Debt load	Cofer & Somers, 2010
Satisfaction with cost	Settle, 2011
Financial assistance (loans, grants, etc.)	Settle, 2011; Fike & Fike, 2008; Cofer & Somers, 1998, 2010; Madgett & Belanger, 2008; Herzog, 2003; McElroy, 2005; Ma & Frempong, 2008
Cost of tuition	Cofer & Somers, 2010

Appendix B: Studies of Canadian College Student Retention

Variables Examined	Significant Association	Limitations	Source/Institution
Gender, date of admission, program choice, academic preparedness, engagement	All five variables were significantly associated with attrition, but only academic preparedness and engagement were of practical significance.	Did not examine many of the social, academic, demographic, and financial factors potentially related to retention	Fisher and Engemann, 2009; Fanshaw College of Applied Arts and Technology
Reasons for taking a program, goal commitment, educational ability, academic/social integration, satisfaction/use of services, student characteristics, labour market conditions	Compared to program completers, non-completers had lower educational goals, were less committed to those goals, had lower high school grades, have different demographic characteristics (were more likely to be disabled, Aboriginal, older, female, married, have dependent children, work during college), and were more likely to be influenced by labour market conditions.	Did not examine many of the social, academic, demographic, and financial factors potentially related to retention.	Sarkar, 1993; Saskatchewan Institute of Applied Science and Technology
Demographic characteristics, academic background, support service needs, attitudes, and educational goals, academic and extracurricular behavior, perceptions, and attitudes subsequent to their enrollment, age, sex, high school grades, and other demographic and academic information	Those retained were significantly more likely to have a better overall high school average and higher marks in grade twelve mathematics. They were also demonstrated higher levels of academic integration. Part-time students were more likely to withdraw.	Small sample, high rate of missing information on some variables.	Kirby and Sharpe, 2001; College of the North Atlantic
High school marks, demographics, and psychosocial factors	Poor high school grades and greater age predicted dropout. However, the factors that predict dropout varied by subgroup. Males with a high school average of less than 80% were more likely to dropout than females, and scored lower than females on a number of psychosocial factors.	No mention of financial factors.	Jorgensen, Ferraro, Fichten, and Havel, 2009; Dawson College

Note. Complete source cited in Reference list.

Appendix C: Concise Theories used to Examine Retention

Theory	**Brief Description of Relevance to Retention**	**Source**
Bourdieu's (1986) social capital theory	PSE is a social process due to transactional nature of interactions between students and faculty, other students, family /friends, the institution, community.	Settle, 2011; Wells 2008
Tinto's (1988; 1993) Student Integration Model	Social and academic integration influences educational outcomes including retention.	Fike & Fike, 2008; Nguyen et al., 2010; Bers, 1988; Robertson & Taylor, 2009; Roberts & Styron, 2010; Madgett & Belanger, 2008; Braxton, Vesper, & Hossler, 1995
Bean & Metzner's (1985; 1990) student attrition model	Incoming student characteristics influence students' interaction with institution; environmental factors and intention may also influence decision to persist.	Fike & Fike, 2008; Bers, 1988; Robertson & Taylor, 2009
Swail, Redd, & Perna's Forces acting on the geometric model of student retention and achievement (2003)	Social, institutional, and cognitive factors influence retention.	Madgett & Belanger, 2008
Astin's (1991; 1984) input-environment-output model	Input factors (e.g. background) and environmental factors (e.g. faculty interaction) must be evaluated when considering output measures (e.g. retention).	Fike & Fike, 2008

Note. Complete source cited in Reference list.

Appendix D: Previous Studies of Retention using Logistic Regression

Definition of Retention	Purpose	Study Design	Sample Description	Results	Limitations	Author and Date Published	Country
Year-to-year enrolment	To develop/ test a theoretical framework to describe retention at 2-year institutions; to develop & compare models for all students, 1st gen students, and continuing gen students.	Quantitative; retrospective; multivariable logistic regression to examine relative effect of 37 factors on retention	n=310 first-time beginning students at 2 year colleges.	Differences existed between models for 1st gen and continuing gen students. All 3 models predicted high proportions of those retained. Having friends attend same institution and having social contact with faculty were strongest predictors. Others incl: age >21, 2+ family members in college, FT attendance, direct entry, satisfaction with cost, financial aid, living on campus.	No mention of variable selection procedures beyond literature review.	Settle, 2011	USA
a)term-to-term retention b) year-to-year retention	To analyze predictors of retention	Quantitative; retrospective; multivariate logistic regression;	N=9200 first time students over a four year period (data collapsed across years). Unclear if this was a 2 or 4 year college.	Term-to-term retention: 65.7%-70.7% across years. Year-to-year retention: 45.8%-49.4% across years. Passing developmental reading and math courses strong predictors of retention.	Missing info on parental level of edu coded as not attended – this may have biased the results.	Fike & Fike, 2008	USA
Term-to-term	To see how background,	Quantitative; retrospective;	n=7507 students at 2-year	57% retention rate. Model accurately predicted 75%	Data are from 1990s, so data	Cofer & Somers,	USA

Definition of Retention	Purpose	Study Design	Sample Description	Results	Limitations	Author and Date Published	Country
retention	achievement, college experiences, price, and debtload influence retention.	Multiple logistic regression.	institutions	of those retained and 72% of attrition.	may be relatively outdated and/or results no longer applicable to students today (in a different economic climate).	2000	
Year-to-year retention	To examine the effects of taking a college orientation course on retention	Quantitative; prospective; Multiple logistic regression.	n=3750 first time students enrolled in fall; unclear if it was 2 or 4 year college	Age, GPA, number of counseling appointments, and enrolment in orientation course were best predictors of retention.	Self-selection bias	Nguyen et al., 2010	USA
Year-to-year retention	To examine the role of social and cultural capital on retention at 2 and 4-year institutions	Quantitative, retrospective; multiple logistic regression.	n=1726 direct entry college and university first-year students with a HS diploma	Higher social and cultural capital linked to higher rates of retention. Retention rates were similar at 2 and 4-year institutions.	Restricted sample does not accurately reflect college demographics	Wells, 2008	USA

Definition	Purpose	Study Design	Sample	Results	Limitations	Author and	Country

of Retention			Description			Date Published	
Year-to-year retention	To identify predictors of retention among incoming students	Quantitative, prospective cohort; forward stepwise logistic regression. Model fit assessed using Log Likelihood.	n=1134 first-time students who were enrolled in the first week aged 18-19; students belonged to the College of Human Environmental Sciences at a university	Several variables predicted retention and accounted for 26.1% of the variance. Several of the covariates were significant but were not strong predictors.	Sample limited to students in a specific field. May not have examined all factors related to retention. Backwards elimination may have been more appropriate method.	Robertson & Taylor, 2009	USA
Not clear, suspect it was year-to-year retention at the same institution	To identify variables that will aid in management of first university experiences.	Quantitative; retrospective, multiple logistic regression	n=4229 students enrolled in their first university experience (included 1st and second year students)	Financial, social systems/habit, academic/cognitive habits, and institutional category predicted retention.	Significant factors may not all be appropriate for intervention; sample limited to universities.	Madgett & Belanger, 2008.	Canada

Note. Complete reference cited in References.

Appendix E: Significant Associations between Covariates – Aggregate Data

	GPA		
Age	Above 50th Percentile	Below 50th Percentile	Total
>20 years	201 47.6% 47.0%	221 52.4% 54.0%	422 100% 50.4%
≤20 years	227 54.7 53.0	188 45.3% 46.0%	415 100% 49.6%
Total	428 51.1% 100%	409 48.9% 100%	837 100% 100%

p=0.041

	First Generation Student Status			
Age	Don't Know	No	Yes	Total
>20 years	26 6.2% 46.4%	296 70.1% 56.3%	100 23.7% 39.4%	422 100% 50.5%
≤20 years	30 7.3% 53.6%	230 55.6% 43.7%	154 37.2% 60.6%	414 100% 49.5%
Total	56 6.7% 100%	526 62.9% 100%	254 30.4% 100%	836 100% 100%

p=0.000

	Passed Grade 12 Math		
Age	No	Yes	Total
>20 years	111 26.9% 34.9%	301 73.1% 56.0%	412 100% 50.2%
≤20 years	207 50.7% 65.1%	201 49.3% 40.0%	408 100% 49.8%
Total	318 38.8% 100%	502 61.2% 100%	820 100% 100%

p=0.000

	High School Average			
Age	<60%	61-80%	81-100%	Total
>20 years	264 63.2% 50.2%	138 33.0% 60.0%	16 3.8% 24.2%	418 100% 50.9%
≤20 years	262 64.9% 49.8%	92 22.8% 40.0%	50 12.4% 75.8%	404 100% 49.2%
Total	526 64.0% 100%	230 28.0% 100%	66 8.0% 100%	822 100% 100%

p=0.000

	Gender		
GPA	Female	Male	Total
Above 50th Percentile	319 74.5% 55.4%	109 25.5% 41.8%	428 100% 51.1%
Below 50th Percentile	257 62.8% 44.6%	152 37.2% 58.2%	409 100% 48.9%
Total	576 68.8% 100%	261 31.2% 100%	837 100% 100%

p=0.000

	First Generation Student Status			
GPA	Don't Know	No	Yes	Total
Above 50th Percentile	19 4.4% 33.9%	276 64.5% 52.5%	133 31.1% 52.4%	428 100% 51.2%
Below 50th Percentile	37 9.1% 66.1%	250 61.3% 47.5%	121 29.7% 47.6%	408 100% 48.8%
Total	56 6.7% 100%	526 62.9% 100%	254 30.4% 100%	836 100% 100%

p=0.028

GPA	Passed Grade 12 Math		Total
	No	Yes	
Above 50th Percentile	132 31.3% 41.5%	290 68.7% 57.8%	422 100% 51.5%
Below 50th Percentile	186 46.7% 58.5%	212 53.3% 42.2%	398 100% 48.5%
Total	328 38.8% 100%	502 61.2% 100%	820 100% 100%

p=0.000

GPA	High School Average			Total
	<60%	61-80%	81-100%	
Above 50th Percentile	239 56.6% 45.4%	155 36.7% 67.4%	28 6.6%	422 100% 51.3%
Below 50th Percentile	287 71.8% 54.6%	75 18.8% 32.6%	42.4%	400 100% 48.7%
Total	526 64.0% 100%	230 28.0% 100%	38 9.5% 57.6%	822 100% 100%

p=0.000

Gender	First Generation			Total
	Yes	No	Don't Know	
Female	189 32.8% 74.4%	353 61.3% 67.1%	34 5.9% 60.7%	576 100% 68.9%
Male	65 25.0% 25.6%	173 66.5% 32.9%	22 8.5% 39.3%	260 100% 31.1%
Total	254 30.4% 100%	526 62.9% 100%	56 6.7% 100%	836 100% 100%

p=0.047

Gender	Passed Grade 12 Math		Total
	No	Yes	
Female	240 42.4% 75.5%	326 57.6% 64.9%	566 100% 69.0%
Male	78 30.7% 24.5%	176 69.3% 35.1%	254 100% 31.0%
Total	318 38.8% 100%	502 61.2% 100%	820 100% 100%

p=0.001

First Generation Student Status	Passed Grade 12 Math		Total
	No	Yes	
Don't Know	30 55.6% 9.4%	24 44.4% 4.8%	54 100% 6.6%
No	179 32.9% 53.5%	347 67.1% 69.1%	517 100% 63.1%
Yes	118 47.4% 37.1%	131 52.6% 26.1%	249 100% 30.4%
Total	318 38.8% 100%	502 61.2% 100%	820 100% 100%

p=0.000

First Generation Student Status	High School Average			Total
	<60%	61-80%	81-100%	
Don't Know	38 70.4% 7.2%	8 14.8% 3.5%	8 14.8% 12.1%	54 100% 6.6%
No	325 62.5% 61.9%	164 31.5% 71.3%	31 6.0% 47.0%	520 100% 63.3%
Yes	162 65.6% 30.9%	58 23.5% 25.2%	27 10.9% 40.9%	247 100% 30.1%
Total	525 64.0% 100%	230 28.0% 100%	66 8.0% 100%	821 100% 100%

p=0.000

Passed Grade 12 Math	High School Average			Total
	<60%	61-80%	81-100%	
No	230 73.7% 44.3%	31 9.9% 13.7%	51 16.4% 79.7%	312 100% 38.5%
yes	290 58.2% 55.8%	195 39.2% 86.3%	13 2.6% 20.3%	498 100% 61.5%
Total	520 64.2% 100%	226 27.9% 100%	64 7.9% 100%	810 100% 100%

p=0.000

Appendix F: Significant Associations between Covariates – College A Data

Gender	GPA		Total
	Above 50th Percentile	Below 50th Percentile	
Female	251 59.2% 71.3%	173 40.8% 59.3%	424 100% 65.8%
Male	101 45.9% 28.7%	119 54.1% 40.8%	220 100% 34.2%
Total	352 54.7% 100%	292 45.3% 100%	644 100% 100%

p=0.001

Grade 12 Pure Math	Hours Spent Working per Week				Total
	1-10	11-19	20+	Not Working	
Did Not Pass	51 27.1% 26.7%	54 28.7% 34.6%	30 16.0% 41.7%	53 28.2% 25.4%	188 100% 29.9%
Passed	140 31.8% 73.3%	102 23.2% 65.4%	42 9.6% 58.3%	156 35.5% 75.6%	440 100% 70.1%
Total	191 30.4% 100%	156 24.8% 100%	72 11.5% 100%	209 33.3% 100%	628 100% 100%

p=0.025

Grade 12 Pure Math	Credit Load		Total
	Full-time	Part-time	
Did Not Pass	127 67.2% 24.8%	62 32.8% 51.7%	189 100% 29.9%
Passed	386 86.9% 75.2%	58 13.1% 48.3%	444 100% 70.1%
Total	513 81.0% 100%	120 19.0% 100%	633 100% 100%

p=0.000

Grade 12 Pure Math	Age		Total
	≤20	>20	
Did Not Pass	76 40.2% 21.7%	113 59.8% 39.9%	189 100% 29.9%
Passed	274 61.7% 78.3%	170 38.3% 60.1%	444 100% 70.1%
Total	350 55.3% 100%	293 44.7% 100%	633 100% 100%

p=0.000

Grade 12 Pure Math	GPA		Total
	Above 50th Percentile	Below 50th Percentile	
Did Not Pass	86 45.5% 24.7%	103 54.5% 36.1%	189 100% 29.9%
Passed	262 59.0% 75.3%	182 41.0% 63.9%	444 100% 70.1%
Total	348 55.0% 100%	285 45.0% 100%	633 100% 100%

p=0.002

Grade 12 Pure Math	Credential			Total
	University Transfer	Upgrading/Non-Parchment	Parchment	
Did Not Pass	8 4.2% 4.0%	67 35.5% 65.7%	114 60.3% 34.7%	189 100% 29.9%
Passed	194 43.7% 96.0%	35 7.9% 34.3%	215 48.4% 65.4%	444 100% 70.1%
Total	202 31.9% 100%	102 16.1% 100%	329 52.0% 100%	633 100% 100%

p=0.000

Hours Spent Studying per Week	Hours Spent Working per Week				
	1-10	11-19	20+	Not Working	Total
0-5	8 22.9% 4.2%	2 5.7% 1.3%	12 34.3% 16.0%	13 37.1% 6.2%	35 100% 5.5%
6-10	57 28.8% 29.5%	50 25.3% 31.9%	18 9.1% 24.0%	73 36.2% 8.1%	214 100% 33.8%
11-15	74 34.6% 38.3%	54 25.2% 34.4%	18 8.4% 24.0%	68 31.8% 32.5%	140 100% 22.15
16-20	43 30.7% 22.3%	38 27.1% 24.2%	21 15.0% 28.0%	38 27.1% 18.2%	47 100% 7.4%
21+	11 23.4% 5.7%	13 28.7% 8.3%	6 12.8% 8.0%	17 36.2% 8.1%	198 100% 31.2%
Total	193 30.4% 1005	157 24.8% 100%	75 11.8% 100%	209 33.0% 100%	634 100% 100%

P=0.002

Appendix G: Significant Associations between Covariates – College B Data

Gender	Dependent Children		Total
	No	Yes	
Female	73	77	150
	48.7%	51.3%	100%
	67.0%	93.9%	78.5%
Male	36	5	41
	87.8%	12.2%	100%
	33.0%	6.1%	21.5%
Total	109	82	191
	57.1%	42.9%	100%
	100%	100%	100%

p=0.000

Gender	GPA		Total
	Above 50th Percentile	Below 50th Percentile	
Female	68	84	152
	44.7%	55.3%	100%
	89.5%	71.8%	78.8%
Male	8	33	41
	19.5%	80.5%	100%
	10.5%	28.2%	21.2%
Total	76	117	193
	39.4%	60.6%	100%
	100%	100%	100%

p=0.003

Gender	Credential			Total
	University Transfer	Upgrading/Non-Parchment	Parchment	
Female	9	56	87	152
	5.9%	36.8%	57.2%	100%
	40.9%	74.7%	90.6%	78.8%
Male	13	19	9	41
	31.7%	46.3%	22.0%	100%
	59.1%	25.3%	9.4%	21.2%
Total	22	75	96	193
	11.4%	38.9%	49.7%	100%
	100%	100%	100%	100%

p=0.000

Aboriginal Status	Grade 12 Pure Math		Total
	Passed	Did Not Pass	
Declared	67	12	79
	84.8%	15.2%	100%
	51.9%	20.7%	42.3%
Not Declared	62	46	108
	57.4%	42.6%	100%
	48.1%	79.3%	57.8%
Total	129	58	187
	69.0%	31.0%	100%
	100%	100%	100%

p=0.000

Dependent Children	Grade 12 Pure Math		Total
	Passed	Did Not Pass	
No	65	41	106
	61.3%	38.7%	100%
	50.8%	71.9%	57.3%
Yes	63	16	79
	79.8%	20.3%	100%
	49.2%	28.1%	42.7%
Total	128	57	185
	69.2%	30.8%	100%
	100%	100%	100%

p=0.007

Dependent Children	Credit Load		Total
	Full-time	Part-time	
No	74	35	109
	67.9%	32.1%	100%
	63.3%	47.3%	57.1%
Yes	43	39	82
	52.4%	47.6%	100%
	36.8%	52.7%	42.9%
Total	117	74	191
	61.3%	38.7%	100%
	100%	100%	100%

p=0.030

Dependent Children	Credential			Total
	University Transfer	Upgrading/Non-Parchment	Parchment	
No	19 17.4% 86.4%	34 31.2% 45.3%	56 51.4% 59.6%	109 100% 57.1%
Yes	3 3.7% 13.6%	41 50.0% 54.7%	38 46.3% 40.4%	82 100% 42.9%
Total	22 11.5% 100%	75 39.3% 100%	94 49.2% 100%	191 100% 100%

p=0.002

Dependent Children	Living Arrangements				Total
	College Residence	Other	Parent's Home	Rental	
No	17 15.6% 42.5%	11 10.1% 40.7%	54 49.5% 79.4%	27 24.8% 48.2%	109 100% 57.1%
Yes	23 28.1% 57.5%	16 19.5% 59.3%	14 17.1% 20.6%	29 35.4% 51.8%	82 100% 42.9%
Total	40 20.9% 100%	27 14.1% 100%	68 35.6% 100%	56 29.3% 100%	191 100% 100%

p=0.000

Grade 12 Pure Math	High School Average			Total
	<60%	61-80%	81-100%	
Did Not Pass	85 67.5% 64.4%	15 11.9% 60.0%	26 20.6% 96.3%	126 100% 68.5%
Passed	47 81.0% 35.6%	10 17.2% 40.0%	1 1.7% 3.7%	58 100% 31.5%
Total	132 71.7% 100%	25 13.6% 100%	27 14.7% 100%	184 100% 100%

p=0.003

Grade 12 Pure Math	Age		Total
	≤20	>20	
Did Not Pass	35 27.1% 56.5%	94 72.9% 75.2%	129 100% 69.0%
Passed	27 46.6% 43.6%	31 53.5% 24.8%	58 100% 31.0%
Total	62 33.2% 100%	125 66.8% 100%	187 100% 100%

p=0.009

Grade 12 Pure Math	Credential			Total
	University Transfer	Upgrading/Non-Parchment	Parchment	
Did Not Pass	9 7.0% 42.9%	67 51.9% 94.4%	53 41.1% 55.8%	129 100% 69.0%
Passed	12 20.7% 57.1%	4 6.9% 5.6%	42 72.4% 44.2%	58 100% 31.0%
Total	21 11.2% 100%	71 38.0% 100%	95 50.8% 100%	187 100% 100%

Appendix H: Significant Associations between Covariates – University Transfer Data

Hours Spent Working /Week	Credit Load		Total
	Full-time	Part-time	
1-10	64 86.5% 33.3%	10 13.5% 31.3%	74 100% 33.0%
11-19	42 82.4% 21.9%	9 17.7% 28.1%	51 100% 22.8%
20+	7 53.9% 3.7%	6 46.2% 18.8%	13 100% 5.8%
Not Working	79 91.9% 41.2%	7 8.1% 21.9%	86 100% 38.4%
Total	192 85.7% 100%	32 14.3% 100%	224 100% 100%

p=0.003

GPA	Credit Load		Total
	Full-time	Part-time	
Above 50th Percentile	103 91.1% 53.7%	10 8.9% 31.3%	113 100% 50.5%
Below 50th Percentile	89 80.2% 46.4%	22 19.8% 68.8%	111 100% 49.6%
Total	192 85.7% 100%	32 14.3% 100%	224 100% 100%

p=0.019

Age	Credit Load		Total
	Full-time	Part-time	
>20	146 89.0% 76.0%	18 11.0% 56.3%	164 100% 73.2%
≤20	46 76.7% 24.0%	14 23.3% 43.8%	60 100% 26.8%
Total	192 85.7% 100%	32 14.3% 100%	224 100% 100%

p=0.019

Appendix I: Significant Associations between Covariates – Upgrading/Non-Parchment Data

Gender	GPA		Total
	Above 50th Percentile	Below 50th Percentile	
Female	60 47.2% 87.0%	67 52.8% 59.3%	127 100% 69.8%
Male	9 37.9% 100%	46 83.6% 40.7%	55 100% 30.2%
Total	69 37.9% 100%	113 62.1% 100%	192 100% 100%

p=0.000

Grade 12 Pure Math	GPA		Total
	Above 50th Percentile	Below 50th Percentile	
Did Not Pass	45 33.6% 66.2%	89 66.4% 84.8%	134 100% 77.5%
Passed	23 59.0% 33.8%	16 41.0% 15.2%	39 100% 22.5%
Total	68 39.3% 100%	105 60.7% 100%	173 100% 100%

p=0.004

Grade 12 Pure Math	Age		Total
	>20	≤20	
Did Not Pass	40 28.9% 65.6%	94 70.2% 83.9%	134 100% 77.5%
Passed	21 53.9% 34.4%	18 46.2% 16.1%	39 100% 22.5%
Total	61 35.3% 100%	112 64.7% 100%	173 100% 100%

p=0.006

Grade 12 Pure Math	Credit Load		Total
	Full-time	Part-time	
Did Not Pass	59 44.0% 69.4%	75 56.0% 85.2%	134 100% 77.5%
Passed	26 66.7% 30.6%	13 33.3% 14.8%	39 100% 22.5%
Total	85 49.1% 100%	88 50.9% 100%	173 100% 100%

p=0.013

GPA	Credit Load		Total
	Full-time	Part-time	
Above 50th Percentile	41 59.4% 46.1%	28 40.6% 30.1%	69 100% 37.9%
Below 50th Percentile	48 42.5% 53.9%	65 57.5% 69.9%	113 100% 62.1%
Total	89 48.9% 100%	93 51.1% 100%	182 100% 100%

p=0.027

Appendix J: Significant Associations between Covariates – Parchment Data

Gender	Grade 12 Pure Math		Total
	Did Not Pass	Passed	
Female	137 44.6% 82.0%	170 55.4% 66.2%	307 100% 72.4%
Male	30 25.6% 18.0%	87 74.4% 33.9%	117 100% 27.6%
Total	167 39.4% 100%	257 60.6% 100%	424 100% 100%

p=0.000

Grade 12 Pure Math	GPA		Total
	Above 50th Percentile	Below 50th Percentile	
Did Not Pass	81 48.5% 33.6%	86 51.5% 47.0%	167 100% 39.4%
Passed	160 62.3% 66.4%	97 37.7% 53.0%	257 100% 60.6%
Total	241 56.8% 100%	183 43.2% 100%	424 100% 100%

p=0.005

Grade 12 Pure Math	Age		Total
	>20	≤20	
Did Not Pass	64 38.3% 34.0%	103 61.7% 43.6%	167 100% 39.4%
Passed	124 48.3% 66.0%	133 51.8% 56.4%	257 100% 60.6%
Total	188 44.3% 100%	236 55.7% 100%	242 100% 100%

p=0.044

GPA	Credit Load		Total
	Full-time	Part-time	
Above 50th Percentile	221 89.8% 61.7%	25 10.2% 34.3%	246 100% 57.1%
Below 50th Percentile	137 74.1% 38.3%	48 26.0% 65.8%	185 100% 42.9%
Total	358 83.1% 100%	73 16.9% 100%	431 100% 100%

p=0.000