Examining the effects of college courses on student outcomes using a joint nearest neighbor matching procedure on a state-wide university system

Nichole D. Smith, MA
Department of Leadership, Policy, and Adult and Higher Education
North Carolina State University
Raleigh, NC 27695
nichole_smith@ncsu.edu

March 2016

Abstract

The purpose of this quasi-experimental study is to examine the effect of online courses on four-year university students' retention and grades in lower-level English and math courses within the University of North Carolina (UNC) system. By using UNC system-wide data, including 15 four-year universities with varying institutional missions, locations, and student populations, the results are highly representative of the four-year university student population. I employ a joint, nearest neighbor matching procedure which utilizes both propensity score caliper matching and Mahalanobis distance matching to reduce selection bias before using ordinary least squares (OLS) regression models to determine effect estimates on both course retention and performance. Matching results show the standardized bias across covariates is significantly decreased and the matched control group is similar to the treatment group, thus significantly reducing selection bias. These results show there is a significant ($p < 0.01$) 5.7 percentage point penalty for course withdrawal and a significant ($p < 0.01$) 0.18 grade point penalty (on a 4-point scale) for online enrollments. The penalties are still significant ($p < 0.01$), but slightly less (3.7 percentage points and 0.18 grade points) for math enrollments and slightly higher (5.9 percentage points and -0.28 grade points) for English enrollments.

Introduction

Approximately 37 percent of undergraduate students took at least one online course in Fall 2013 (Allen & Seaman, 2014). For students enrolled in the University of North Carolina (UNC) system in the 2013-14 academic year, this proportion was even higher (39 percent) (UNC-GA, 2015). From popular annotated bibliographies (e.g., Russell, 1999) to U.S. Department of Education meta-analyses (e.g., Means, Toyama, Murphy, Bakia, & Jones, 2009) and from scholarly literature reviews (e.g., Hart, 2012; Schlosser & Anderson, 1994; Tallent-Runnels et al., 2006) to media articles (e.g., Carey, 2015; Kolowich, 2015), it seems a majority of the population has embraced the growth of education in the virtual realm.

Historically, Russell's book "No Significant Difference Phenomenon" (1999), an annotated bibliography, has strongly influenced the conversation around the quality of online courses. Unfortunately, it is often overlooked that Russell's compilation of analyses spans a wide spectrum of course-types (K-12 to workforce development), modes of delivery (hybrid courses to online trainings), methodologies (descriptive statistics to randomized experiments), and time (1928 to 2013). It does not immediately follow from this collection that there is no significant difference in course outcomes between online and face-to-face four-year university students. In fact, only 35 research studies on four-year university online student outcomes were identified between a 2009 meta-analysis funded by the U.S. Department of Education (Means et al., 2009), along with a 2010 revision (Means, Toyama, Murphy, & Jones, 2010), 2013 follow-up (Lack, 2013), and subsequent 2015 literature review from Ithaka S+R (Wu, 2015). An even smaller subset of these studies, approximately six, utilize experimental or quasi-experimental methods to control for selection effects.

A quasi-experimental methodology is necessary for evaluating the effect of online courses since students choose to take a course online versus face-to-face, or self-select. The factors that drive students to select into an online course are likely correlated with student outcomes. This is problematic since the self-selection process also naturally sorts students such that the treated group (students taking a course online) is very different from the untreated group (students taking the same course face-to-face) on many demographic and descriptive characteristics. Because these characteristics are correlated with the factors driving self-selection, and the factors driving self-selection are correlated with the outcomes of interest, any evaluation that does not account for the selection process into online courses (e.g., t-tests, chi-square tests, or simple linear regression) is biased (Khandker, Koolwal, & Samad, 2010).

The purpose of this study is to examine the effect of online courses on four-year university students' retention and grades in lower-level English and math courses within the University of North Carolina (UNC) system. I employ a joint, nearest neighbor matching procedure which utilizes both propensity score caliper matching and Mahalanobis distance matching to reduce selection bias before using ordinary least squares (OLS) regression models to determine effect estimates on both course retention and performance. This methodology offers three significant additions to the literature: 1) By focusing on lower-level English and math courses, I offer a four-year university student point of comparison to Xu and Jaggars' (2011a) community college student effect estimates. 2) The UNC system-wide data includes 15 four-year universities with varying institutional missions, locations, and student populations. This makes the results highly representative of other four-year university student populations. 3) Using a joint matching procedure involving both propensity scores and a Mahalanobis distance calculation generates a control group that is very similar to the treatment group on a host of covariates. Because these groups are more similar than the full population of face-to-face and online students, the bias in the effect estimates are significantly reduced.

In order to specify the models so that the effect estimates are as reliable as possible, it is important to review the literature and ensure the proper control variables are included. For example, matching models should include covariates that affect both treatment and outcomes, and outcome models should control for factors that could impact estimated effects. The below theoretical framework identifies what characteristics impact both selection into online courses and the outcomes of interest. The literature review provides context for how other studies have addressed this framework and the existing estimates for the impact of online courses on student outcomes. Following the review, I describe the UNC system dataset, explain the sampling procedure, and relate the available data to the concepts explained in the literature. Next, I explain the joint nearest neighbor matching method, the various methodologies involved, and the empirical models used. This explanation is followed by the matching results and effect estimates for course retention and performance. Finally, I summarize the results and explain what impact they have on my research moving forward.

Theoretical Framework

The crux of this study is the generation of the propensity scores and calculation of a Mahalanobis distance. The covariates used in any matching procedure, particularly propensity scores, should be factors the literature shows are linked to both participation in the observed treatment and the desired outcome (Barth, Guo, & McCrae, 2008). Consequently, my conceptual framework covers both what is known about selection into treatment—enrollment into an online course—and what is known about the outcomes of interest—student retention and student performance. Unfortunately, the increase of online enrollments has occurred so rapidly that educational theory and research have been unable to keep up. For this reason, I utilize my own compilation of literature surrounding online students and online student outcomes to build up a framework for selection that is then utilized juxtapose with Rovai's (2003) model for online student retention to inform methodological decisions.

Modeling Selection into Online Courses

There is no theoretical model for the selection process into an online course, but the literature is unanimous about one thing: online students are different from face-to-face students on a host of variables. Tallent-Runnels' et al. (2006) literature review cites the average age of online learners between 29 and 35 years; this age range falls outside of the traditional college student age range (18-24 years) (NCES, n.d.). Many other studies corroborate this finding (CCRC, 2013; Diaz & Bontenbal, 2001; Dupin-Bryant, 2004; Eduventures, 2008; Sussman & Dutter, 2010). As students age out of the traditional range, students' lives become multi-dimensional; they are more likely to be financially independent, employed, married, have children, etc. The decision for a traditional student to enroll in an online course over a face-to-face course may be a simple cost-benefit analysis on the number of available seats in a face-to-face section or the distance of the walk to class, but that decision is much more complicated for a non-traditional student. This student population is likely to be weighing the benefits of working from home over the commute to a college campus; or the benefits of completing coursework after the kids have gone to bed over attending class during dinner time; or the benefits of being able to work a full-time schedule over having to take off work multiple times a week for class.

My review of the literature corroborates this idea that the decision to enroll in an online course is based on a cost-benefit analysis. Online student populations are more likely to be employed full-time (CCRC, 2013; Dutton et al., 2002; Sussman & Dutter, 2010; McLaren, 2004), be married (CCRC, 2013; Sussman & Dutter, 2010), and have children (Dutton et al., 2002). In addition to these external responsibilities, online learners are reported to be different types of learners as well; they are reported to be more self-directed (Diaz & Bontenbal, 2001; Dutton et al., 2002; Liu et al., 2007; Stewart et al., 2010), motivated (Dutton et al., 2002), willing to collaborate (Diaz & Bontenbal, 2001; Dutton et al., 2002), and experienced with computers (Dutton et al., 2002).

Modeling Course Retention and Course Performance

Rovai's (2003) of online student retention is designed to capture the phenomenon of attrition for the online student population. His framework combines elements from Tinto's theory of student retention (Tinto, 1975, 1987, 1993) with elements from Bean and Metzner's student attrition model (1985) along with relevant research for online student populations in an effort to better capture online student experiences. Because Rovai's (2003) model (figure 1) only includes literature through 2002, I interject my own review of the literature to identify factors influencing online student retention. There is not a performance model for online students, but throughout my review of the literature, I realized that the characteristics impacting student retention also significantly affected performance. For this reason, I adopt Rovai's (2003) model for my performance outcome model as well. The following breakdown of relevant constructs offer evidence that the factors are important for both models.

Rovai's (2003) model groups the variables that impact online student retention into four major categories: a) student characteristics prior to college admission, b) student skills prior to college admission, c) external factors after college admission, and d) internal factors after college admission. While these groups correspond to a student's pathway into and through college, institutions do not capture information about their students with this process in mind. So, in order to ensure my quantitative models adequately control for these variables, I have reorganized them into the following six categories: a) student background, b) student characteristics, c) student outcomes, d) institutional characteristics, e) student interaction with institution characteristics, and f) external factors.

Student Background. The student background characteristics are taken from both Tinto's theory of student retention (1997, 1987, 1993) and Bean and Metzner's student attrition model (1985) for face-to-face students, but are confirmed to affect online student retention as well. Many studies find that gender significantly impacts online course performance (Chyung, 2001; Gunn, McSporran, Macleod, & French, 2003; Price, 2006; Rovai & Baker, 2005; Sullivan, 2001; Taplin & Jegede, 2001) and that females appear to complete online courses at a higher rates (Aragon & Johnson, 2008; Kupczynski, Mundy, Goswami, & Meling, 2014; Rovai, 2001). Rovai (2003) reports minority students feel isolated in online courses (Murguia, Padilla, & Pavel, 1991), and ethnicity has also been reported to significantly affect online course performance (Newell, 2007). Furthermore, older, more advanced students are reportedly more likely to complete (Didia & Hasnat, 1998; Dupin-Bryant, 2004; Levy, 2007; Moore et al., 2002; UNC-GA, 2013; Wojciechowski & Palmer, 2005) and pass (Wang & Newlin, 2002; Willging & Johnson, 2004) their online courses.

Student Outcomes. Rovai's (2003) model suggests students who have a history of performing well in other courses or have prior experience with online courses will persist. Recent research corroborates Rovai's (2003) model, suggesting higher performing students are more likely to persist (Aragon & Johnson, 2008; Dupin-Bryant, 2004; Levy, 2007; Moore, Bartkovich, Fetzner, & Ison, 2002; Morris, Wu, & Finnegan, 2005) and are more successful in distance learning courses (Figlio, Rush, & Yin, 2010; Schlosser & Anderson, 1994).

Additionally, the more online courses students take, a proxy for computer literacy, the more likely they are to complete (Aragon & Johnson, 2008; Boston, Ice, & Gibson, 2011).

Student Characteristics. In addition to tangible outcomes, Rovai (2003) includes student characteristics in the model separately from student outcomes since various researchers find that online coursework requires an additional set of skills over face-to-face coursework (Cole, 2000; Rowntree, 1995). More recent studies report that students who lack time management skills drop out (Nash, 2005; Moore et al., 2002) or are lower performing in online courses. Rovai (2003) cites Sherry (1996), Moore and Kearsley (1996), Grow (1996), and Tucker (2000) in saying a more self-directed learning style makes a student better prepared for success in online courses. Furthermore, my review found that internal factors such as lack of motivation, teaching style, goal commitment, self-efficacy, and satisfaction are barriers to online student retention (Hart, 2012; Moore et al., 2002; Nash, 2005; Street, 2010).

Student Plus Institution Characteristics. The student is not directly responsible for all characteristics that affect retention. The student's relationship with the institution can greatly impact retention and performance, which is why Rovai (2003) includes Tinto's (1975, 1987, 1993) factors for academic and social integration, goal and institutional commitment, and learning community along with Bean and Metzner's (1985) program fit factor.

Institutional Characteristics. To better capture the institutional characteristics outside of a student's control, Rovai's (2003) model includes Workman and Stenard's (1996) needs of distance education learners program clarity, self-esteem, interpersonal relationships, and accessibility to services. In agreement, the literature shows lack of support (Nash, 2005), academic subject (Jaggars, 2012; UNC-GA, 2013), and grade level (UNC-GA, 2013) significantly affect online student performance.

External Factors Finally, Rovai (2003) includes environmental or external factors from Bean and Metzner's (1985) attrition model such as finances, employment, familial responsibilities, etc. Literature confirms the impact of such elements on online student retention. Family or childcare, finances, and job needs are often cited as barriers to course completion for online students (e.g., Martinez, 2003; Moore et al., 2002).

Literature Review

Online Course Outcomes without Selection Controls. Descriptively, course withdrawal rates have been estimated to be between 10 to 20 percentage points higher for online four-year university students (Carr, 2000). However, because we know online student populations to be vastly different from campus-based students, any studies that do not use some sort of quasi-experimental methods to account for the selection process into online courses are severely biased. After controlling for various characteristics, skills, and factors outlined above, many find online courses have a negative effect on student retention (Frydenberg, 2007; McLaren, 2004; Patterson & McFadden, 2009). However, one community college study found no significant difference between online and face-to-face course retention (Enriquez, 2010), and another found that online students were actually more likely to complete the course than their face-to-face peers (Aragon & Johnson, 2008). The only study for four-year undergraduate students finds immediate course retention is lower for online students as evidenced by a chi-square test (McLaren, 2004), but we know this kind of analysis is likely biased.

It is important to keep in mind that students are only assigned a grade if they complete a course. Based on the literature above for online student course retention, a larger portion of the online students in the following studies already dropped the course before they could receive a grade of pass or fail. Because students report low performance as a common reason for course withdrawal (Dunwoody & Frank, 1995; Reed, 1981), higher dropout rates are likely to inflate the average course grade for online courses in observational studies that do no control for selection or sufficient covariates. This may explain why two observational studies have found online student course performance to be higher than face-to-face student performance; these studies compared their populations using t-tests (Reasons, Valadares & Slavkin, 2005; Rieger, 2002). It may also explain why many other studies find no significant difference when using observational methods varying from pre and posttest analyses (Ary & Brune, 2011; Johnson, 2002), linear models (Daymont & Blau, 2008; Friday, Friday-Stroud, Green, & Hill, 2006; Lu et al., 2003), and t-tests or chi-squared tests (CDHE, 2012; Enriquez, 2010; McLaren, 2004; Plumb & LaMere, 2011; Summers et al., 2005; Wagner, Garippo & Lovaas, 2011).

Even with the potential overestimation of online student performance, a handful of observational studies find course performance to be lower for online students across observational methods (Brown & Liedholm, 2002; Carter, 2012; Ferguson & Tryjankowski, 2009; Jones & Long, 2013; Parsons-Pollard, Lacks, & Grant, 2008; Samman, Omar, Belmasrour, & Alijani, 2013; Schwartz, 2012; Weber & Lennon, 2007). In summary, the only two studies for four-year university students find lower performance for online students and are likely underestimating this negative effect.

Online Course Outcomes with Selection Controls. Studies that lowered selection bias (Xu & Jaggars, 2011a, 2011b, 2013) found that performance was higher for face-to-face students than online students across three quasi-experimental studies on community college students using a particular sample selection, fixed effects, and propensity score matching. The first study (Xu & Jaggars, 2011b), involving Washington State Community and Technical College students, attempted to control for selection bias by limiting their sample to only students who had ever taken an online course. This method creates a counterfactual from online course-takers (between 25 and 45 percent of the student population depending on the year), rather than their face-to-face only peers. This eliminates students from the analyses who have a low propensity to take an online course. They also constructed hierarchical linear models to account for the students nested within institutions; these models controlled for various student, course, and institution level variables. The results showed face-to-face student course retention and performance was higher than that of online students across models.

The second study (Xu & Jaggars, 2013) utilizes a fixed effects model to estimate the effects of online courses on student retention and performance for the same population of Washington State Community and Technical College students and finds the same results. What is unique about this study is the way they address selection through their fixed effects model and robustness checks. Their base model includes student level demographic and academic preparedness covariates along with semester level covariates, such as the number of credits taken that semester. They then insert a variable for individual fixed effects and create the course-demeaned data on each outcome (the average of the fixed effects equation across course observations subtracted from the original equation). This removes any omitted variable bias due to course-invariant, student level factors. To address time and course level biases, they include semester and academic subject fixed effects. Their results show that face-to-face student course retention and performance was higher than that of online students.

Lastly, Xu and Jaggars (2011a) used propensity score matching before performing logistic regression to account for selection bias. They provide a table of descriptive statistics to illustrate the differences between online and face-to-face course takers and determine that “online course takers are substantially different from classroom-based students; in particular, online takers tend to have stronger academic preparation" (p. 362). They include student level demographic covariates (e.g., age, gender, ethnicity, etc.), student level preparation covariates (e.g., dual enrollment, financial aid, enrollment status, etc.), and school level covariates (e.g., instructional expenditures, location in a rural area, percentage of minority students, etc.) in their multi-level propensity score equation; they use the same covariates in their empirical models. To create the counterfactual, Xu & Jaggars (2011a) use an exact matching method rather than a weighting method. They identify balanced covariates by providing probability densities for the groups before and after matching; this serves to justify the selection bias has been reduced. In the end, they report face-to-face student course retention and performance to be higher than that of online students.

Methodology

This quasi-experimental method uses nearest neighbor matching to identify a sample of control students against whom I can compare the treatment group. A student's nearest neighbor is defined using "Mahalanobis metric matching within calipers defined by the estimated propensity score" (Rosenbaum & Rubin, 1985). Compared to other methods that rely on the Mahalanobis metric or propensities alone, this process has been deemed superior for balancing both covariates and propensity scores (D'Agastino Jr., 1998; Rosenbaum & Rubin, 1985; Sekhon, 2007). Producing balance on the propensity score is important for treatments such as online course enrollment since the distribution of propensities for the control group is largely skewed toward zero.

This joint process maintains the property from the individual methods of being equal-percent bias reducing (EPBR); this means "the percent reduction in the (mean) bias of each of the covariates is the same" (Rubin, 1976). Other methods may lower the bias for some covariates while raising the bias for others, but an EPBR method reduces the bias proportionally for all covariates. However, because the matching method will not completely remove the bias between samples, I will employ ordinary least squares (OLS) regression on the treatment group and matched control groups to estimate the average treatment effects on the treated (ATETs) for retention and performance. Covariates used for the creation of the propensity scores are selected based on their impact on both treatment and outcomes. These same variables will be used in the calculation of the Mahalanobis distance between students. They are also included in the regression models for each outcome, along with additional covariates specifically affecting the outcome. In addition to these covariates, the final regression models will include fixed effects to control for institution level unobservables.

Because these analyses are conducted over two outcomes (course withdrawal and course grade) and two course subjects (both combined and individually), I completed six total matching procedures. The propensity scores are calculated three separate times, once for the combined enrollments, once for the English enrollments, and once for the math enrollments. Then, because course completion is required in order to receive a valid course grade, each of these sets of propensity scores are then used in two joint matching procedures, once for all enrollments and once for course completers (only students with a final grade).

Institutional Dataset and Sample

This study utilizes transcript data for over 4.2 million course enrollments from the 15 four-year institutions within the University of North Carolina (UNC) system for the 2012-13 and 2013-14 academic years[1]. Institutions within the UNC system span all geographic regions of the

[1] Each academic year is comprised of two semesters—fall and spring. Summer courses have been excluded since the short length of a summer semester impacts the implementation of both online and face-to-face courses.

state and vary drastically in size, demographics, and course offerings. For example, two of the 15 institutions enroll over 30,000 students, while two enroll fewer than 5,000. The system includes four historically Black institutions and one historically American Indian institution. The percentage of students at an institution taking at least one online course ranges from 5 percent to 77 percent. In order to maintain confidentiality, the institutions have been randomly ordered and assigned a number 1 through 15.

My analyses are conducted at the enrollment level since my outcomes are retention and performance within a course. An individual enrollment represents a student within a course section within a course within an institution. This nested structure makes the sampling procedure very technical since decisions were needed at each level to identify the final sample of enrollments. For example, when an undergraduate student is enrolled in a graduate course, the student’s level is recognized as undergraduate, but the enrollment is considered to be at the graduate level. To address this particular misalignment, I removed courses that were considered “graduate level” and I removed course enrollments completed by “graduate students”. This left only undergraduate students with enrollments in undergraduate courses. The nested structure also adds a level of complexity to a course section’s mode of delivery. To remove as much variability as possible within my treatment variable, I excluded any course sections with enrollments of conflicting modes of delivery.

Overall, I utilized a four-stage process to identify my analytic sample (depicted in table 1). Stage one removed ineligible courses. I removed any course sections with a hybrid or face-to-face distance education mode of delivery. Further explanation is provided in the treatment section below. I also removed any graduate or upper-level undergraduate courses. This analysis is limited to lower-level courses in order to eliminate confounding factors that are introduced as students move into upper-level, major-specific courses. In stage two, I only included enrollments completed by an undergraduate student. Furthermore, if a student presented multiple enrollments within the same course during the available academic years, only the first enrollment was included. Stage three used the outcome data to further narrow the sample to credit-earning enrollments (i.e., course audits are excluded) and enrollments for students who completed at least one course during the semester. For example, if a student withdrew from the institution during a semester, all of his/her course enrollments for that semester are excluded. These exclusions are further explained within the outcome section below. Finally, stage four identified courses with no more than a ratio of two online enrollments per every face-to-face enrollment between the 2012-13 or 2013-14 academic year. These enrollments were not required to occur within the same semester. After implementing these four stages, the final analytic sample consisted of 155,792 online enrollments and 765,714 face-to-face enrollments. Specific to this paper, the sample was further narrowed down to English and math enrollments. This serves as a point of comparison for Xu and Jaggars' (2011a) community college study, but also serves as a starting point of my research plan to expand these analyses to other course subjects across the system.

Derived Variables

The institutional data is structured for the purposes of reporting and accountability and is not immediately structured for analytical purposes. Creating outcome measures, treatment variables, and covariate information is time-intensive and technical. This section lays out the necessary technical information for understanding how variables of interest were captured, and outlines the context and relevance of each in modeling online course effects.

Outcomes: Online Course Retention and Performance. The two outcomes of interest for this study are course retention and course performance. Course retention will be measured through the variable *withdrawal*, which receives a value of 1 if the student withdrew from the course after the census date but before the end of the semester and a value of 0 if the student completed the course and received a legitimate grade. Students are not penalized for withdrawing before the census date nor are they identified in the dataset, but withdrawing after the census date can affect students' financial aid, enrollment status, or graduation requirements. It is generally in a student's best interest to remain enrolled after the census date unless remaining enrolled will have a significant, negative impact on a student's grade point average (GPA). The dataset does not identify whether the student was passing or failing at the time of withdrawal, but for the reasons above, course withdrawal is considered a negative course outcome. I have removed any students who withdrew from all their courses in a semester since complete withdrawal from the institution is likely due to external reasons unrelated to course effects.

Second, I model course performance using an ordinary least squares equation on the subsamples of overall course completions, math completions, and English completions. Course performance is a continuous variable measured by the variable *grade*. The institutional datasets include the number of quality points earned in a course and the number of credit hours assigned to the course. The variable *grade* divides the number of quality points by the number of credit hours to create a final course grade for the student based on a four-point scale. Any grade assigned as "pass" rather than a numeric value was set to a grade of 4.0. Any enrollments assigned zero credit hours (i.e., the course was being audited) were excluded, as were any grades above a 5.0 since this is too high for the data to be reliable.

Treatment: Online Course Enrollment. I chose to begin my analyses with 2012-13 because the UNC system systematically changed the way institutions designate online enrollments. This change took place during the 2011-12 academic year, so 2012-13 is the first year with reliable data using the new definitions, which states an online course is, “A course with instruction delivered through the Internet or World Wide Web that does not have dedicated physical space and meeting times. The course may have physical space and meeting times during testing or other examination periods.” The UNC system differentiates online courses with a face-to-face portion by designating them as hybrid courses and indicating whether the percentage of time spent in a physical meeting space is above or below 50%. Because the range for the amount of time spent interacting face-to-face in these hybrid courses is so wide and could significantly impact course outcomes, I have excluded all hybrid sections from the sample. Site-based distance education classes have also been excluded since these are face-to-face sections but are not offered on campus and may introduce confounding effects. The treatment variable *course_online* is a course-section-by-semester level variable that is equal to 1 if the course section was completed online and 0 if the course section was conducted face-to-face. The treatment group is considered to be every course enrollment for which the variable *course_online* is equal to 1.

The covariates are created at the semester level for each student. For this reason, if a student appears in the treatment group multiples times within the same semester, these enrollments are weighted such that each student-by-semester observation is treated with a weight of 1. For example, if one student took 3 treatment courses online within the same semester, each course enrollment for that student is given a frequency weight of 1/3. If one student took two treatment courses online but in different semesters, these course observations are given a weight of 1 since each observation will be unique per the semester-level covariates. Any multiple sample course face-to-face enrollments within the same semester will be weighted similarly for pre-matching analyses, and weighted again separately for the post-matching analyses.

Covariates. The covariates were created to best fit Rovai's (2003) model described in the literature review. This section overviews each piece of the re-structured model and explains how the covariates capture the individual constructs or serve as a proxy. In some cases, the institutional data is not robust enough to capture a particular construct. Table 2 provides the construct associated with each covariate.

Due to large amounts of missingness for variables such as income and high school GPA within the institutional data, many otherwise continuous or dichotomous variables were re-coded and analyzed with a "missing" category. For example, the variable *hsgpa_pctl*, introduced as the percentile for high school GPA below, has five categories 0, 1, 2, 3 (the four percentiles calculated based on the entire sample) and 99, a separate category for anyone with a missing high school GPA. Because the covariates are specific to particular models, table 2 also provides an indication of which model (i.e., propensity score, Mahalanobis distance, or outcome) each variable was included in.

Student Background. This construct captures relevant, innate student characteristics—age, ethnicity, and gender. Each of these are uniquely captured in the dataset through the continuous variable *age*, categorical variable *race* which captures the nine IPEDS racial categories, and indicator variable *male*.

Student Outcomes. Student outcomes both prior to and after college admission play a significant role in Rovai's (2003) model. Both the high school and college academic outcomes are well represented in the institutional data through high school GPA, captured in a continuous variable *hsgpa,* and a categorical variable *hsgpa_pctl* for the high school GPA percentile, SAT scores, captured in a continuous variable *sat,* and a categorical variable *sat_pctl* for SAT percentile, and cumulative college GPA, captured in the continuous and categorical variables *uncgpa* and *uncgpa_pctl*.

Unfortunately, the institutional data does not have any high school level transcript information, so I am unable to capture computer literacy or computer-based interactions prior to the available academic years. However, the Department of Commerce shows correlations between digital literacy and demographics such as age, race, academic level, urbanicity, and income (Digital Nation, 2011). Each of these variables is represented in the dataset, and together they proxy a student's level of computer literacy. It has already been explained that age and race are represented within the dataset. The academic year of a student is captured in the ordinal variable *acad_lvl*. The data contains the zip code for a student's permanent residence if the residence is within the state of North Carolina. Based on zip code, an indicator, *rural*, for whether the zip code is considered urban or rural (as according the The Rural Center[2]). A student's income is represented in the continuous variable *income*, but a categorical variable *income_pctl* is also in the dataset. Similarly, the institutional data does not capture any academic behaviors such as study habits and absenteeism in college. Nevertheless, they are naturally correlated with a student's overall college GPA, captured in the variable *uncgpa* introduced above.

Student Characteristics. Individual student characteristics such as goal commitment, self-esteem, satisfaction and stress levels, and learning styles are difficult constructs to measure and only identifiable by the individual student. As such, these factors are not captured within the institutional dataset and are not controlled for within these analyses. This is a common limitation

[2] http://www.ncruralcenter.org/rural-data

among studies utilizing administrative data, but the trade-off is that the UNC data provide an extremely large sample across which many individual personalities are captured.

Student Plus Institution Characteristics. Institution and program fit are going to have a significant impact on the outcomes of interest since a student who is not invested in the school or the program are going to be more disengaged than their peers. While the dataset does not capture students' personal opinions about fit, we can proxy this fit by controlling for how long a student has persisted at the institution (through the variable *acad_lvl*) and whether the student's major (*cip_major* captured by a 2-digit CIP code) is the same as or related to the same course subject as the current course enrollment. Relation is measured through an indicator of whether the CIP code is identified as a STEM field (science, technology, engineering or mathematics), *stem*.

Institutional Characteristics Along with the student level predictors, Rovai's (2003) model also explains how institution, program, and course level characteristics can impact retention and performance. Each institution and academic subject area (i.e., English and math) will be controlled for through dummy variables. This controls for factors such as services available to the student and course availability. Unfortunately, one factor that cannot be controlled for is teaching styles. One glaring limitation of the existing institutional data is that it has not been merged with instructor level variables. This prevents me from controlling for any instructor level factors that may impact the outcomes of interest. Fixed effects are not even possible since I cannot identify whether three sections were taught by three different instructors or were all taught by the same instructor.

External Factors Finally, external factors such as finances, family and employee responsibilities, outside encouragement, and life events influence the outcomes of interest. Finances are captured by the income variables mentioned above, as well as an indicator for whether the student received a Pell grant, *pell*. External responsibilities are proxied through the indicator variable for whether a student is considered independent for financial aid purposes, *independent*, a student's household income, a student's age, and the number of credit hours enrolled that semester, *num_credhr*. Life events are handled by the sampling procedure which excludes course enrollments for students who withdrew from all courses during a particular semester. Outside encouragement is proxied through the number of family member currently enrolled in college, *num_college*, since one would expect a student with other family members in college to have more support for their education-related activities than someone who does not have other family members in college. Unfortunately, other common proxies for this construct such as parental education or marital status are not available.

Joint Nearest Neighbor Matching Procedure

Nearest neighbor matching procedures are used to identify one control subject as a match for each treatment subject based on a specific set of criteria for the covariates of interest. Choosing this matched subsample is meant to decrease bias in the effect estimates by balancing the covariate means such that they are not statistically significantly different between the treatment and control groups. Oftentimes a nearest neighbor is defined as the person who is closest based on a multi-dimensional distance metric (e.g., a Mahalanobis distance) or a propensity score. Matching using a multi-dimensional distance metric is efficient for creating balance across each of the covariate means, but even when the propensity score is included in the distance measure, the balance on the propensity score is not always sufficient. On the other hand, matching on the propensity score alone may not balance the covariate means as closely as a Mahalanobis distance match. Rosenbaum and Rubin (1985) identified a joint procedure using both propensity score calipers and the Mahalanobis distance that balances the covariate means as sufficiently as the Mahalanobis distance matching and balances the propensity scores with the same effect of propensity score matching. Their joint procedure is implemented as follows:

1. Calculate a propensity score for all sample members.
2. Randomly order the treatment and control groups.
3. Select the first/next treatment student, called stu_t_j, and let p_j be student j's propensity score.
4. Identify all control students with a propensity score within $p_j \pm \alpha$, where α is a pre-determined caliper. Call these potential matches stu_c_i for each student i. (If no stu_c_i are identified, stu_t_j is dropped, and the process returns to step 3.)
5. Calculate a Mahalnobis distance between stu_t_j and stu_c_i for all i, called mh_{ji}.
6. Select stu_c_i with the smallest mh_{ji}, called stu_c_{ji}.
7. Remove stu_c_{ji} from the list of control students.
8. Repeat steps 3-7 for all treatment students.
9. Generate a list of all students, stu_c_{ji}. This is the matched control group.

I implement a modified version of the above procedure to my sample of course enrollments. In an attempt to find the best match, I match with replacement and later weight the duplicate enrollments. So, step 7 is dropped. Then, because this sample size is very large, I often have the ability to implement stricter criteria than what is identified by Rosenbaum and Rubin's original step 4. In order to identify the best matches possible (and decrease the processing time of my procedures), I have added additional criteria to step four. First, I identified all control students within the same institution and course subject as stu_t_j. Of these students, I chose the first of the following four subgroups that was not empty to be used in step 5:

a. All control students with a propensity score within $p_j \pm \alpha$ who were enrolled in the same course and are of the same gender.
b. All control students with a propensity score within $p_j \pm \alpha$ who were enrolled in the same course.
c. All control students with a propensity score within $p_j \pm \alpha$ who are of the same gender.
d. All control students with a propensity score within $p_j \pm \alpha$.

If each subgroup listed above is empty, the treatment student is removed from the analysis. In order to implement the joint procedure, it is first necessary to understand how propensity scores and Mahalanobis distances are calculated. The following sections outline these methods.

Propensity Scores. The term propensity score originated with Rosenbaum and Rubin (1983) and is defined to be "the conditional probability of assignment to a particular treatment given a vector of observed covariates." Propensity scores take all of the variables we know to affect selection into treatment (enrollment in an online course) and condense the complicated process of selection down into one value for each observation through the following logistic regression equation:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_i X_i \ ,$$

where p is the probability that a student selects into an online course, and X_i is the vector of observed covariates for student i referred to by Rosenbaum and Rubin (1983). The covariates for my propensity score model are indicated by the corresponding column of table 2.

In order to identify the potential control matches in step 4 above, a caliper of α= "20% of a standard deviation of the overall propensity score distribution" was chosen to minimize the number of potential matches, while ensuring sufficient coverage. This calculated to α=0.020, 0.026, and 0.018 for the overall model, English-only model, and math-only model respectively.

With these calipers, the full set of treatment enrollments had, on average, 18 potential control student matches identified within the caliper. As separate samples, the English and math enrollments had 28 and 12 potential control student matches identified on average, respectively.

Mahalanobis Distance. The Mahalanobis distance was then calculated between the treatment student, *stu_t$_j$*, and each of the potential control matches. The Mahalanobis metric is a multi-dimensional metric by which the distance between two vectors of data can be measured. For reference, the distance between two-dimensional vectors simplifies down to the well-known formula for the distance between two points:

$$d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

In higher dimensions, the Mahalanobis distance between vectors **u** and **v** becomes:

$$d(\mathbf{u}, \mathbf{v}) = (\mathbf{u} - \mathbf{v})^T \mathbf{C}^{-1} (\mathbf{u} - \mathbf{v})$$

where **C** is the covariance matrix of the covariates from the set of control students identified in step 4 of the matching process. For the purposes of this study, the two vectors of interest are the vectors of covariates for *stu_t$_j$* and *stu_cj$_i$*; the specific covariates included in the distance measure are indicated by the corresponding column of table 2. All covariates were normalized before the distance was calculated. This prevents covariates with a high range, such as income, from dominating the distance measure when compared to those with small ranges (e.g., dummy variables). After normalization, the furthest apart two students can be, based on an individual covariate, is 1. The average distance between all enrollments in the overall sample was 2.29. In the separated sample of full enrollments for English and math, the average distances were 2.12 and 2.44 respectively.

One constraint of implementing this procedure is that this metric requires the covariates to be continuous, dichotomous, or ordinal so all categorical covariates needed transformation prior to the Mahalanobis distance calculation. I implemented two types of transformations. The first turned a categorical variable of *n* categories into *n* dichotomous variables. This was applied to *race, pell, independent, and cip_major* since, for example, the "distance" between an Asian student and a black student is not conceptually different from the "distance" between a white student and a black student. The second type of transformation was applied to *income, sat, hsgpa*, and *uncgpa* Each of these variables is a naturally continuous variable that was originally transformed into a categorical variable to account for missingness. For the distance metric, these variables were reverted back to continuous variables. Consider the covariate *income*. This transformation allows two students with income values to be compared on their exact income values. If two students had missing values, their values were set to the covariate mean (a normalized value of 0.5); this allows all other covariates to determine the students' overall distance. Finally, when one student has an income value and one student's is missing, they are set to a distance of 1 on that covariate; this sets the students who cannot be compared as far apart as possible, erring on the side of the two students being a bad match. See figure 2 for examples of the effects of this transformation on the Mahalanobis distance between two students. In addition to these transformations, the variable *stem* was added to the Malahanobis distance calculation. Its inclusion allowed for a closer measure of student major. Otherwise any two majors would be equally as far apart as any other two.

Empirical Models

As discussed in the literature review, there are many non-cognitive factors that go into the decision to enroll in an online course, none of which are measurable through institutional data. While I have discussed ways in which the available data will serve as proxies for these unobserved factors, the matching procedure will not remove all bias from the counterfactual. In an effort to further reduce bias before comparing outcomes of interest for the two groups, I will generate doubly robust average treatment effects through OLS models that include the covariates used in matching, along with the additional proxy for outside encouragement, *num_college*. See the full list of variable associated with the empirical models in the corresponding column of table 2.

The empirical model for the outcome of course withdrawal, *withdraw,* was conducted on the entire sample. Because *withdraw* is a binomial outcome, logistic regression would be the expected method of analysis. However, further investigation showed that the high number of categorical variables affected the fit of a logistic regression model. Based on a lower AIC and BIC when compared to logistic regression, I chose to calculate the treatment effect on course withdraw using OLS. A model for the *grade* outcome was conducted on the set of course completers (those who did not withdraw).

Results

Overall

Descriptively, the full population of enrollments (both English and math combined) in online courses is quite different from the face-to-face population. See table 3a for a list of pre-matching descriptive statistics for both the full sample of enrollments and the sample of course completions. These differences align with what the literature predicts. For example, the online student population is much less likely to be first year students and much more likely to be fourth year students. They are also more likely to be older, file FAFSA as an independent, and be female. They also show lower high school GPAs and SAT scores. Prior to matching, 10 covariates (21 values overall) exhibited standardized bias values higher than 20%. The standardized bias for the propensity score was 137%.

The matching procedures are almost identical for the sample of all enrollments and the sample of completions. The following results and referenced figures discuss the matching results for the full sample of enrollments, but the tables referenced provide the corresponding values for the sample of course completions. Overall, the matching procedure was unable to match 422 enrollments for the withdrawal analysis and an additional 9 enrollments for the grade analysis, leaving the final samples with 14,839 and 13,227 enrollments respectively. After matching, only the covariate for age and the propensity score had standardized differences greater than 20%. See table 4a for the list of post-matching descriptive statistics. Figure 3a depicts the decrease in standardized bias for relevant covariates for the full sample of enrollments. The standardized bias decreased considerably for all covariate values; the decreases range from 19 to 100 percent. The values for first, third, and fifth year undergraduate levels, the highest income quartile, and lowest high school GPA quartile all received a 90% decrease in standardized bias or higher. Some values for missing categories also received a 90% decrease or higher, but because these values encompass so many different scenarios, they are impossible to interpret.

The post-matching regression models show slight increases in the penalty for online courses as compared to the population models, but the effect is significantly negative at the one percent level for both outcomes. See table 5. According to these models, the ATET is a 5.7 percentage point increase in the probability of course withdrawal ($p < 0.01$) and a 0.18 grade point penalty on a 4-point scale for online course enrollments ($p < 0.01$).

Math

Prior to matching, the set of math enrollments look fairly similar to the overall population, with one exception being that the set of math students were generally more likely to be first year undergraduates and the online enrollments in particular were more likely to be first year undergraduates than the overall online population. The same set of covariates exhibited a standardized bias of over 20%. See table 3b for post-matching values across both the full sample of enrollments and sample of completions; see figure 3b for a depiction of the difference in standardized bias before and after matching for relevant covariates. After matching, the final samples were 7,923 and 6,087 for the withdrawal and grade regression models respectively. The matched models (table 5) show a significant 3.6 percentage point increase in the probability of withdrawing from an online math course ($p < 0.01$) and a 0.11 grade point penalty ($p < 0.01$).

English

The population of English enrollments, like the math enrollments, is very similar to the population. These students, opposite of the math students, are less likely to be first year students and more likely to be third and fourth year undergraduates. Again, the same set of covariates shows standardized biases over 20%. See table 3c for post-matching values across both the full sample of enrollments and sample of completions; see figure 3c for a depiction of the difference in standardized bias before and after matching for relevant covariates. After matching, the final samples were 5,131 and 4,611 for the withdrawal and grade regression models respectively, and the ATETs show a significant 5.9 percentage point increase in the probability of withdraw ($p < 0.01$) and a 0.28 grade point penalty ($p < 0.01$).

Conclusion and Discussion

The overall negative impact online enrollments have on course retention and performance are consistent with the literature, but an all-encompassing model such as this does not have enough nuance to directly impact policy. Breaking down these analyses by course subject allows institution and department leaders to look at specific programs and courses to decide how best to structure them. However, the outcomes for the English and math enrollments are opposite of what is expected. Comparable studies do not control for selection or employ different matching techniques, but this does not explain the significantly higher negative impact online English enrollments have over math enrollments both in the raw data and in the matched sample. This contradicts both the literature and intuition. There are a few places in this analysis that will need further investigation in order to identify why these effects would be reversed. For example, the exclusion of hybrid courses, upper level courses, and the exclusion of online courses that do not have a face-to-face counterpart may subset the English enrollments down to a group that, indeed does experience a more negative effect than math enrollments. It may also be that institutions already have policies in place that offer more assistance to online math enrollments, so the negative effect is already being addressed. Or, due to the fact that the matched group is significantly less biased and the control group is significantly more comparable to the group of English enrollments, these analyses show an effect previously undetermined. It could be that with multitudes of help documents and external mathematical resources, it is much easier for online math students to find examples and extra materials to help them through the class than it would be for an English student.

In summary, the matching procedure is successful in identifying a control group so that the covariate means are comparable and the outcomes can be more reliably compared. There is a significant negative penalty associated with both retention rates and performance for online enrollments. These negative penalties are still present when analyzing math subjects and English subjects individually, though further investigation is warranted to identify what may be causing counter-intuitive results. Next steps will involve expanding the sample to include a more extensive set of course subjects, calibrating the matching and empirical models to ensure I have the best fitting models, and conducting robustness checks. These checks will involve both adding in hybrid courses and upper level courses, and adding an empirical model with course pass/fail rates as an outcome.

References

Allen, I. E., & Seaman, J. (2014). Grade Change: Tracking online education in the United States, 2013. *Babson Survey Research Group and Quahog Research Group, LLC.*

Aragon, S. R., & Johnson, E. S. (2008). Factors influencing completion and noncompletion of community college online courses. *The Amer. Jrnl. of Distance Education*, *22*(3), 146-158.

Ary, E. J., & Brune, C. W. (2011). A comparison of student learning outcomes in traditional and online personal finance courses. *Journal of Online Learning and Teaching*, *7*(4), 465-474.

Barth, R. P., Guo, S., & McCrae, J. S. (2008). Propensity score matching strategies for evaluating the success of child and family service programs. *Research on Social Work Practice*, *18*(3), 212-222.

Bean, J., & Metzner, B. (1985). A conceptual model of nontraditional undergraduate student attrition. Review of Educational Research, 55, 485–650.

Boston, W. E., Ice, P., & Gibson, A. M. (2011). Comprehensive assessment of student retention in online learning environments. *Online Journal of Distance Learning Administration*, *14*(4).

Brown, B. W., & Liedholm, C. E. (2002). Can web courses replace the classroom in principles of microeconomics? *American Economic Review*, 444-448.

Carey, K. (2015, March 5). Here's What Will Truly Change Higher Education: Online Degrees That Are Seen as Official. *New York Times.* Retrieved from www.nytimes.com

Carr, S. (2000). As distance education comes of age, the challenge is keeping the students. The Chronicle of Higher Education, 46(23), A39-A41.

Carter, L. W. (2012). *Determining if instructional delivery model differences exist in remedial English* (Doctoral dissertation, Liberty University).

Chyung, S. Y. (2001). Systematic and systemic approaches to reducing attrition rates in online higher education. The American Journal of Distance Education, 15(3), 36‒49.

Cole, R. A. (Ed.) (2000). Issues in web-based pedagogy: a critical primer. London: Greenwood Press.

Colorado Department of Higher Education (CDHE). (2012). Online Versus Traditional Learning: A Comparison Study of Colorado Community College Science Classes. *WICHE Cooperative for Education Technologies.* Retrieved from wcet.wiche.edu

Community College Research Center (CCRC). (2013). *What We Know About Online Course Outcomes.* New York, NY: Columbia University.

D'Agastino Jr, R. B. (1998). Propensity score methods for bias reduction in the comparison of a treatment group to a nonrandomized control group.*Stat Med*, *17*(19), 2265-2281.

Daymont, T., & Blau, G. (2008). Student performance in online and traditional sections of an undergraduate management course. *Journal of Behavioral and Applied Management*, *9*(3), 275-294.

Diaz, D. P. & Bontenbal, K. F. (2001). Learner preferences: Developing a learner-centered environment in the online or mediated classroom. Ed at a Distance, 15(8)

Didia, D., & Hasnat, B. (1998). The determinants of performance in the university introductory finance course. Financial Practice and Education, 8(1), 102–107.

Digital Nation: Expanding Internet Usage, NTIA Research Preview (Feb. 2011) at 5,http://www.ntia.doc.gov/reports/2011/NTIA_Internet_Use_Report_February_2011.pdf.

Dunwoody, P. T., & Frank, M. L. (1995). Why students withdraw from classes. *The Journal of*

Psychology, *129*(5), 553-558.

Dupin-Bryant, P. A. (2004). Pre-entry variables related to retention in online distance education. *American Journal of Distance Education*, 18(4), 199-206. Retrieved from doi.org/10.1207/s15389286ajde1804_2

Dutton, J., Dutton, M., & Perry, J. (2002). How do online students differ from lecture students? Journal of Asynchronous Learning Networks, 6(1).

Eduventures, LLC. (2008). The big picture: Market size, forecasts and differentiation in online higher education. Boston, MA.

Enriquez, A. (2010). Assessing the Effectiveness of Synchronous Content Delivery in an Online Introductory Circuits Analysis Course. *EDUCATING ENGINEERS IN THE WILD, WILD WEST*, 48.

Ferguson, J., & Tryjankowski, A. M. (2009). Online versus face-to-face learning: looking at modes of instruction in Master's-level courses. *Journal of Further and Higher Education*, *33*(3), 219-228.

Figlio, D. N., Rush, M., & Yin, L. (2010). Is it live or is it internet? Experimental estimates of the effects of online instruction on student learning (NBER Working Paper No. 16089). Cambridge, MA: National Bureau of Economic Research.

Friday, E., Friday-Stroud, S. S., Green, A. L., & Hill, A. Y. (2006). A multi-semester comparison of student performance between multiple traditional and online sections of two management courses. *Journal of Behavioral and Applied Management*, *8*(1), 66-81.

Frydenberg, J. (2007). Persistence in university continuing education online classes. *The International Review of Research in Open and Distributed Learning*, *8*(3).

Grow, G. O. (1996). Teaching learners to be self-directed. Adult Education Quarterly, 41(3), 125–149.

Gunn, C., McSporran, M., Macleod, H., & French, S. (2003). Dominant or different? Gender issues in computer supported learning. Journal of Asynchronous Learning Networks, 7(1), 14−30.

Hart, C. (2012). Factors Associated With Student Persistence in an Online Program of Study: A Review of the Literature. *Journal of Interactive Online Learning*, *11*(1).

Jaggars, S. S. (2012, April). Beyond flexibility: Why students choose online courses in community college. Paper presented at the American Educational Research Association Annual Meeting, Vancouver, Canada.

Johnson, M. (2002). Introductory Biology" Online": Assessing Outcomes of Two Student Populations. *Journal of College Science Teaching*, *31*(5), 312-17.

Jones, S. J., & Long, V. M. (2013). Learning Equity between Online and On-Site Mathematics Courses. *Journal of Online Learning & Teaching*, *9*(1).

Khandker, S. R., Koolwal, G. B., & Samad, H. A. (2010). *Handbook on impact evaluation: quantitative methods and practices*. World Bank Publications.

Kolowich, S. (2015, April 10). How 'Elite' Universities Are Using Online Education. *The Chronicle of Higher Education.* Retrieved from http://chronicle.com

Kupczynski, L., Mundy, M. A., Goswami, J., & Meling, V. (2014). Cooperative learning in distance learning: a mixed methods study.

Lack, K. A. (2013). Current status of research on online learning in postsecondary education. *Ithaka S+ R*, 2013.

Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers and*

Education, 48, 185-204.

Liu, S., Gomez, J., Khan, B., & Yen, C. (2007). Toward a learner-oriented community college online course dropout framework. International Journal on E-Learning, 6(4), 519–542.

Lu, J., Yu, C. S., & Liu, C. (2003). Learning style, learning patterns, and learning performance in a WebCT-based MIS course. *Information & Management*, *40*(6), 497-507.

Martinez, M. (2003, July 14). High attrition rate in e-learning: Challenges, predictors, and solutions. The eLearning Developers Journal, 1-7.

McLaren, C. H. (2004). A comparison of student persistence and performance in online and classroom business statistics experiences. *Decision Sciences Journal of Innovative Education*, *2*(1), 1-10.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies. *US Department of Education*.

Means, B., Toyama, Y., Murphy, R., & Jones, K. (2010). Evaluation of evidence-based practices in online learning: a meta-analysis and review of online learning studies. US Department of Education Office of Planning, Evaluation, and Policy Development Policy and Program Studies Service. Revised September 2010.*USA Today.[Online]*. *Available: http://www. usatoday. com/story/news/nation/2013/07/11/moocs-top-colleges-and-universities/2509883*.

Moore, K., Bartkovich, J., Fetzner, M., & Ison, S. (2002). Success in Cyberspace: Student Retention in Online Courses. AIR 2002 Forum Paper.

Moore, M. G., & Kearsley, G. (1996). Distance education: a systems view. Belmont, CA: Wadsworth Publishing Company.

Morris, L. V., Wu, S. S., & Finnegan, C. L. (2005). Predicting retention in online general education courses. *The American Journal of Distance Education*, *19*(1), 23-36.

Murguia, E., Padilla, R. V., & Pavel, M. (1991). Ethnicity and the concept of social integration in Tinto's model of institutional departure. Journal of College Student Development, 32, 433–439.

Nash, R. (2005, Winter). Course completion rates among distance learners: Identifying possible methods to improve retention. Online Journal of Distance Learning Administration, 8(4).

National Center for Education Statistics (NCES). (n.d.). *Nontraditional Undergraduates: Definitions and Data.* Retrieved from http://nces.ed.gov/pubs/web/97578e.asp.

Newell, C. C. (2007). Learner characteristics as predictors of online course completion among nontraditional technical college students (Unpublished doctoral dissertation). University of Georgia, Athens, GA.

Parsons-Pollard, N., Lacks, R. D., & Grant, P. H. (2008). A comparative assessment of student learning outcomes in large online and traditional campus-based introduction to criminal justice courses. *Criminal Justice Studies*, *21*(3), 239-251.

Patterson, B., & McFadden, C. (2009). Attrition in online and campus degree programs. Online Journal of Distance Learning Administration, 12(2).

Plumb, C., & LaMere, B. (2011). Comparing student learning in a required electrical engineering undergraduate course: Traditional, face-to-face vs. online. In *Annual Conference of the International Network for Engineering Education and Research. August* (pp. 21-26).

Price, L. (2006). Gender differences and similarities in online courses: Challenging stereotypical views of women. Journal of Computer Assisted Learning, 22(5), 349–359.

Reasons, S. G., Valadares, K., & Slavkin, M. (2005). Questioning the hybrid model: Student outcomes in different course formats. *Journal of Asynchronous Learning Networks*, *9*(1), 83-94.

Reed, J. G. (1981). Dropping a college course: Factors influencing students' withdrawal decisions. *Journal of Educational Psychology*, *73*(3), 376.

Rieger, M. (2002). Distance Education versus Classroom Instruction in Horticulture—An Introduction to Fruit Crops—Case Study. *HortTechnology*, *12*(3), 513-515.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*(1), 33-38.

Rubin, D. B. (1976). Multivariate matching methods that are equal percent bias reducing, I: Some examples. *Biometrics*, 109-120.

Rovai, A. P. (2001). Building classroom community at a distance: a case study. Educational Technology Research and Development Journal, 49(4), 35–50.

Rovai, A. P. (2003). In search of higher persistence rates in distance education online programs. *The Internet and Higher Education*, *6*(1), 1-16.

Rovai, A. P., & Baker, J. D. (2005). Gender differences in online learning: Sense of community, perceived learning, and interpersonal interactions. Quarterly Review of Distance Education, 6(1), 31−44.

Rowntree, D. (1995). Teaching and learning online. A correspondence education for the 21st century? British Journal of Educational Technology, 26(3), 205–215.

Russell, T. L. (1999). *The no significant difference phenomenon: A comparative research annotated bibliography on technology for distance education: As reported in 355 research reports, summaries and papers*. North Carolina State University.

Samman, E., Omar, A., Belmasrour, R., & Alijani, G. D. S. (2013). Strategic Plan for Enhancing Online Learning. *Information Systems Education Journal*, *11*(2), 36.

Schlosser, C. A., & Anderson, M. L. (1994). Distance education: review of the literature. Washington, DC: Association for Educational Communications and Technology.

Sekhon, J. S. (2007). Alternative balance metrics for bias reduction in matching methods for causal inference. *Survey Research Center, University of California, Berkeley*.

Sherry, L. (1996). Issues in distance learning. International Journal of Educational Telecommunications, 1(4), 337–365.

Schwartz, D. A. (2012). Effectiveness of learning in online versus on-campus accounting classes: A comparative analysis. *Journal of Research in Innovative Teaching, 5 (1)*, 63-77.

Stewart, C., Bachman, C., & Johnson, R. (2010). Students' characteristics and motivation orientations for online and traditional degree programs. Journal of Online Learning and Teaching, 6(2), 367−379.

Street, H. D. (2010). Factors influencing a learner's decision to drop-out or persist in higher education distance learning. *Online Journal of Distance Learning Administration*, *13*(4).

Sullivan, P. (2001). Gender differences and the online classroom: Male and female college students evaluate their experiences. Community College Journal of Research and Practice, 25(10), 805−818.

Summers, J. J., Waigandt, A., & Whittaker, T. A. (2005). A comparison of student achievement and satisfaction in an online versus a traditional face-to-face statistics class. *Innovative Higher Education*, *29*(3), 233-250.

Sussman, S., & Dutter, L. (2010). Comparing student learning outcomes in face-to-face and online course delivery. *Online Journal of Distance Learning Administration*, *8*(4), 19-30.

Tallent-Runnels, M. K., Thomas, J. A., Lan, W. Y., Cooper, S., Ahern, T. C., Shaw, S. M., & Liu, X. (2006). Teaching courses online: A review of the research. *Review of educational research*, *76*(1), 93-135.

Taplin, M., & Jegede, O. (2001). Gender differences in factors influencing achievement of distance education students. Open Learning, 16(2), 133−154.

Tinto, V. (1975). Dropout from higher education: a theoretical synthesis of recent research. Review of Educational Research, 45(1), 89–125.

Tinto, V. (1987). Leaving college. Chicago: University of Chicago Press.

Tinto, V. (1993). Leaving college: rethinking the causes and cures of student attrition. Chicago: University of Chicago Press.

Tucker, S. Y. (2000). Assessing the effectiveness of distance education versus traditional on-campus education. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

University of North Carolina General Administration (UNC-GA). (2013). *The University of North Carolina Evaluation of E-Learning Across the System.* Chapel Hill, NC: UNC.

University of North Carolina General Administration (UNC-GA). (2015). *The University of North Carolina Enrollment Report: Fall 2014.* Chapel Hill, NC: UNC.

Wagner, S. C., Garippo, S. J., & Lovaas, P. (2011). A longitudinal comparison of online versus traditional instruction. *MERLOT Journal of Online Learning and Teaching*, *7*(1), 68-73.

Wang, A. Y., & Newlin, M. H. (2002). Predictors of performance in the virtual classroom. THE Journal, 29(10), 21−25.

Weber, J. M., & Lennon, R. (2007). Multi-Course Comparison of Traditional versus Web-Based Course Delivery Systems. *Journal of Educators Online*, *4*(2), n2.

Willging, P. A., & Johnson, S. D. (2004). Factors that influence students' decision to dropout of online courses. *Journal of Asynchronous Learning Networks,* 13(3), 115−127.

Wojciechowski, A., & Palmer, L. B. (2005). Individual student characteristics: can any be predictors of success in online classes? *Online Journal of Distance Learning Administration*, 8(2), 13.

Workman, J. J., & Stenard, R. A. (1996). Student support services for distance learners. DEOSNEWS, 6(3)

Wu, D. (2015). Online Learning in Postsecondary Education: A Review of the Empirical Literature. *Ithaka S+R, 2015.*

Xu, D., & Jaggars, S. S. (2011a). The effectiveness of distance education across Virginia's Community Colleges: Evidence from introductory college-level math and English courses. *Educational Evaluation and Policy Analysis*, *33*(3), 360-377.

Xu, D., & Jaggars, S. S. (2011b). Online and Hybrid Course Enrollment and Performance in Washington State Community and Technical Colleges. CCRC Working Paper No. 31. *Community College Research Center, Columbia University*.

Xu, D., & Jaggars, S. S. (2013). Adaptability to Online Learning: Differences across Types of Students and Academic Subject Areas. CCRC Working Paper No. 54. *Community College Research Center, Columbia University*.

Table 1. Four-stage sampling process.

Population of fall and spring enrollments at 4-year institutions			
Enrollments	4,092,486	Course Sections	95,442
Students	309,295	Courses	39,854

STAGE ONE

Remove hybrid and face-to-face distance education sections					
Control		**Treatment**		**Total**	
Enrollments	3,634,438	Enrollments	349,474	Enrollments	3,983,912
Students	272,996	Students	124,804	Students	303,696
Course Sections	83,064	Course Sections	9,753	Course Sections	92,024
Courses	36,923	Courses	5,625	Courses	39,242

Remove graduate and upper-level undergraduate courses					
Control		**Treatment**		**Total**	
Enrollments	3,093,183	Enrollments	262,233	Enrollments	3,355,416
Students	230,715	Students	99,086	Students	247,350
Course Sections	55,894	Course Sections	5,723	Course Sections	60,963
Courses	23,542	Courses	3,091	Courses	24,354

STAGE TWO

Keep only undergraduate student enrollments					
Control		**Treatment**		**Total**	
Enrollments	3,080,473	Enrollments	260,217	Enrollments	3,340,690
Students	225,308	Students	97,819	Students	241,092
Course Sections	55,832	Course Sections	5,703	Course Sections	60,883
Courses	23,502	Courses	3,084	Courses	24,313

Keep first of multiple enrollments					
Control		**Treatment**		**Total**	
Enrollments	2,832,781	Enrollments	249,987	Enrollments	3,082,768
Students	225,259	Students	95,791	Students	241,092
Course Sections	55,415	Course Sections	5,684	Course Sections	60,452
Courses	23,502	Courses	3,080	Courses	24,313

STAGE THREE

Drop semester enrollments for students who withdrew from school

Control		Treatment		Total	
Enrollments	2,735,276	Enrollments	246,586	Enrollments	2,981,862
Students	224,821	Students	94,552	Students	240,721
Course Sections	51,048	Course Sections	5,608	Course Sections	56,018
Courses	23,252	Courses	3,057	Courses	24,053

Drop non-credit enrollments

Control		Treatment		Total	
Enrollments	2,702,882	Enrollments	239,960	Enrollments	2,942,842
Students	222,054	Students	92,407	Students	237,036
Course Sections	51,002	Course Sections	5,592	Course Sections	55,958
Courses	23,233	Courses	3,053	Courses	24,033

STAGE FOUR

Keep courses with a ratio of no more than two online enrollments per every face-to-face (Final analytic enrollment sample)

Control		Treatment		Total	
Enrollments	765,714	Enrollments	155,792	Enrollments	921,506
Students	175,028	Students	71,924	Students	195,237
Course Sections	8,394	Course Sections	3,630	Course Sections	11,495
Courses	1,929	Courses	1,929	Courses	1,929

Table 2. Covariate descriptions, constructs, and corresponding model inclusion.

Level	Variable	Variable Description	Relevant Constructs	Variable Type	Propensity Score	Mahalanobis Distance	Empirical Model
Student by Semester by Institution							
	acad_lvl	Student's Academic Level	Student Outcomes Student Plus Institution Characteristics	categorical	X		X
	cip_major	Student's Major (CIP Code)	Student Plus Institution Characteristics	categorical	X	X	X
	course_subject	Course Subject of Enrollment	Instituitonal Characteristics	categorical	X	X	X
	stem	Student's Major is in a STEM Field	Student Plus Institution Characteristics			X	
	num_credhr	Number of Credit Hours Enrolled	External Factors	continuous	X	X	X
	pell	Pell Recipient	External Factors	categorical	X	X	X
Student by Semester							
	income	Student's Household Income	External Factors Student Outcomes	continuous		X	
	income_pctl	Student's Household Income (Percentile)	External Factors Student Outcomes	categorical	X		X
	independent	Student's Dependency Status	External Factors	categorical	X	X	X
	num_college	Number of People from Household in College	External Factors	categorical			X
	rural	Urbanicity of Permanent Residence	Student Outcomes	categorical	X	X	X
	uncgpa	Student's UNC Cumulative GPA	Student Outcomes Student Plus Institution Characteristics	continuous		X	
	uncgpa_pctl	Student's UNC Cumulative GPA (Percentile)	Student Outcomes Student Plus Institution Characteristics	categorical	X		X
Student							
	age	Student's Age as of September 1, 2012	Student Background Student Outcomes	continuous	X	X	X
	hsgpa	Student's High School GPA	Student Outcomes	continuous		X	
	hsgpa_pctl	Student's High School GPA (Percentile)	Student Outcomes	categorical	X		X
	male	Student's Gender	Student Background	dichotomous	X		X
	race	Student's Race	Student Background Student Outcomes	categorical	X	X	X
	sat	Student's SAT Score	Student Outcomes	continuous		X	
	sat_pctl	Student's SAT Score (Percentile)	Student Outcomes	categorical	X		X
Enrollment							
	inst	Institution of Enrollment	Instituitonal Characteristics	categorical	X	X	X
	time	Semester	Instituitonal Characteristics	categorical	X	X	X

Table 3a. Pre-matching descriptive statistics and standardized differences for relevant covariates by overall sample.

			Pre-matching Overall - Withdraw				Pre-matching Overall - Grade			
Level	Variable	Value	Treated Mean	Untreated Mean	Standardized Difference		Treated Mean	Untreated Mean	Standardized Difference	
Student by semester by institution	Propensity score		0.34	0.06	1.37	*	0.35	0.06	1.39	*
	Number of credit hours		12.97	14.87	-0.58	*	12.86	14.87	-0.61	*
	Received Pell Grant									
		No	7.50	10.60	-0.11		7.70	10.70	-0.11	
		Yes	12.00	13.10	-0.03		11.70	13.10	-0.04	
		Missing	80.50	76.30	0.10		80.60	76.20	0.11	
	Academic level (undergraduate)									
		First year	9.90	55.20	-1.11	*	9.50	55.50	-1.13	*
		Second year	21.70	25.10	-0.08		21.40	24.90	-0.08	
		Third year	30.60	12.40	0.46	*	30.40	12.10	0.46	*
		Fourth year	33.70	6.10	0.74	*	34.30	6.10	0.75	*
		Fifth year	0.40	0.20	0.03		0.40	0.20	0.03	
		Unknown	3.60	1.00	0.17		4.00	1.10	0.19	
	Course subject									
		English	56.10	41.50	0.30	*	56.70	42.20	0.29	*
		Math	43.90	58.50	-0.30	*	43.30	57.80	-0.29	*
Student by semester	Permanent Residence									
		Urban	53.70	53.50	0.01		53.50	53.30	0.00	
		Rural	39.60	33.30	0.13		39.70	33.30	0.13	
		Unknown	6.70	13.30	-0.22	*	6.80	13.40	-0.22	*
	Dependency status									
		Dependent	10.90	22.00	-0.30	*	10.80	22.20	-0.31	*
		Independent	8.60	1.60	0.32	*	8.60	1.60	0.32	*
		Missing	80.50	76.30	0.10		80.60	76.20	0.11	
	Income percentile									
		First percentile	5.40	4.20	0.06		5.30	4.10	0.06	
		Second percentile	5.10	4.80	0.02		5.00	4.80	0.01	
		Third percentile	3.70	5.00	-0.07		3.70	5.10	-0.07	
		Fourth percentile	4.80	9.40	-0.18		4.90	9.50	-0.18	
		Missing	81.00	76.70	0.11		81.10	76.60	0.11	
	UNC GPA percentile									
		First percentile	5.70	4.20	0.07		5.70	4.00	0.08	
		Second percentile	4.80	3.50	0.07		4.50	3.50	0.06	
		Third percentile	4.40	3.20	0.06		4.30	3.10	0.06	
		Fourth percentile	3.80	2.40	0.08		4.00	2.40	0.09	
		Missing	81.30	86.70	-0.15		81.50	86.90	-0.15	

Level	Variable	Value	Pre-matching Overall - Withdraw Treated Mean	Pre-matching Overall - Withdraw Untreated Mean	Pre-matching Overall - Withdraw Standardized Difference		Pre-matching Overall - Grade Treated Mean	Pre-matching Overall - Grade Untreated Mean	Pre-matching Overall - Grade Standardized Difference	
Student	Age		25.25	19.22	0.87	*	25.28	19.16	0.89	*
	Gender	Male	0.33	0.46	-0.26	*	0.33	0.46	-0.25	*
	High school GPA percentile									
		First percentile	17.40	21.30	-0.10		16.70	20.80	-0.11	
		Second percentile	14.80	22.90	-0.21	*	14.80	23.00	-0.21	*
		Third percentile	14.90	20.50	-0.15		15.00	20.80	-0.15	
		Fourth percentile	12.20	21.20	-0.25	*	12.50	21.70	-0.25	*
		Missing	40.70	14.10	0.63	*	40.90	13.70	0.64	*
	SAT score percentile									
		First percentile	12.70	19.40	-0.18		12.30	19.40	-0.20	
		Second percentile	11.40	20.00	-0.24	*	11.30	19.90	-0.24	*
		Third percentile	12.00	17.90	-0.17		12.00	18.10	-0.17	
		Fourth percentile	12.10	19.60	-0.21	*	12.10	19.90	-0.21	*
		Missing	51.90	23.10	0.62	*	52.40	22.70	0.64	*
Enrollment (student by semester by institution)	Semester									
		Fall 2012	24.20	29.00	-0.11		24.20	29.10	-0.11	
		Spring 2013	27.00	22.50	0.11		27.60	22.50	0.12	
		Fall 2013	22.10	27.50	-0.13		21.90	27.60	-0.13	
		Spring 2014	26.70	20.90	0.14		26.40	20.80	0.13	
	Institution									
		1	7.10	7.70	-0.02		7.40	7.70	-0.01	
		2	19.90	25.80	-0.14		20.80	25.80	-0.12	
		3	4.00	12.20	-0.30	*	3.70	11.80	-0.31	*
		4	2.90	1.60	0.09		3.00	1.60	0.10	
		5	1.00	0.20	0.11		0.90	0.20	0.10	
		6	2.60	0.20	0.21	*	2.70	0.20	0.21	*
		7	19.70	17.40	0.06		19.80	17.60	0.06	
		8	10.50	10.90	-0.02		9.20	10.90	-0.06	
		9	9.90	5.50	0.17		9.60	5.50	0.16	
		10	3.00	3.90	-0.05		3.10	4.10	-0.05	
		11	9.40	3.10	0.26	*	9.40	3.10	0.26	*
		12	1.40	2.70	-0.09		1.50	2.70	-0.09	
		13	1.20	0.70	0.05		1.30	0.70	0.06	
		14	7.50	8.10	-0.03		7.60	8.10	-0.02	

* Indicates the standardized bias is greater than 20%

Table 3b. Pre-matching descriptive statistics and standardized differences for relevant covariates by math sample.

			Pre-matching Separate Math - Withdraw				Pre-matching Separate Math - Grade			
Level	Variable	Value	Treated Mean	Untreated Mean	Standardized Difference		Treated Mean	Untreated Mean	Standardized Difference	
Student by semester by institution	Propensity score		0.29	0.05	1.27	*	0.30	0.05	1.28	*
	Number of credit hours		12.75	14.85	-0.62	*	12.65	14.84	-0.64	*
	Received Pell Grant									
		No	6.80	10.80	-0.14		6.90	10.90	-0.14	
		Yes	11.80	12.90	-0.03		11.60	12.90	-0.04	
		Missing	81.40	76.40	0.12		81.40	76.20	0.13	
	Academic level (undergraduate)									
		First year	13.70	59.70	-1.09	*	13.50	60.30	-1.11	*
		Second year	26.30	24.50	0.04		26.30	24.30	0.05	
		Third year	28.60	10.40	0.47	*	28.10	10.10	0.47	*
		Fourth year	26.60	4.00	0.66	*	26.90	3.90	0.67	*
		Fifth year	0.60	0.30	0.05		0.60	0.30	0.04	
		Unknown	4.20	1.10	0.20		4.60	1.10	0.21	*
Student by semester	Permanent Residence									
		Urban	57.00	54.10	0.06		56.60	53.80	0.06	
		Rural	36.00	32.60	0.07		36.20	32.60	0.08	
		Unknown	7.00	13.40	-0.21	*	7.20	13.60	-0.21	*
	Dependency status									
		Dependent	10.80	22.00	-0.31	*	10.90	22.30	-0.31	*
		Independent	7.80	1.60	0.29	*	7.60	1.50	0.30	*
		Missing	81.40	76.40	0.12		81.40	76.20	0.13	
	Income percentile									
		First percentile	4.80	4.00	0.04		4.80	4.00	0.04	
		Second percentile	4.80	4.70	0.01		4.70	4.70	0.00	
		Third percentile	3.90	5.00	-0.05		3.90	5.00	-0.05	
		Fourth percentile	4.50	9.40	-0.20		4.60	9.60	-0.20	
		Missing	82.00	76.90	0.13		82.00	76.70	0.13	
	UNC GPA percentile									
		First percentile	4.40	4.30	0.01		4.40	4.10	0.01	
		Second percentile	4.60	3.50	0.06		4.30	3.40	0.04	
		Third percentile	3.70	3.00	0.04		3.60	3.00	0.04	
		Fourth percentile	2.60	2.20	0.03		2.60	2.20	0.03	
		Missing	84.60	87.00	-0.07		85.10	87.30	-0.06	

Level	Variable	Value	Pre-matching Separate Math - Withdraw Treated Mean	Untreated Mean	Standardized Difference		Pre-matching Separate Math - Grade Treated Mean	Untreated Mean	Standardized Difference	
Student	Age		25.72	19.09	0.94	*	25.67	19.00	0.95	*
	Gender	Male	0.35	0.48	-0.27	*	0.35	0.48	-0.26	*
	High school GPA percentile									
		First percentile	15.60	21.80	-0.16		15.30	21.20	-0.15	
		Second percentile	13.10	22.60	-0.25	*	13.40	22.70	-0.24	*
		Third percentile	13.90	19.40	-0.15		13.70	19.80	-0.16	
		Fourth percentile	10.40	21.20	-0.30	*	10.60	21.90	-0.31	*
		Missing	47.00	14.90	0.74	*	47.00	14.40	0.76	*
	SAT score percentile									
		First percentile	13.00	20.50	-0.20	*	13.00	20.50	-0.20	*
		Second percentile	10.80	20.90	-0.28	*	10.80	20.80	-0.28	*
		Third percentile	10.30	17.40	-0.21	*	10.20	17.60	-0.22	*
		Fourth percentile	10.50	18.40	-0.23	*	10.50	18.70	-0.24	*
		Missing	55.40	22.80	0.71	*	55.50	22.40	0.72	*
Enrollment (student by semester by institution)	Semester									
		Fall 2012	21.30	30.70	-0.22	*	21.20	30.80	-0.22	*
		Spring 2013	28.50	21.00	0.18		28.90	21.00	0.18	
		Fall 2013	20.70	29.00	-0.19		20.50	28.90	-0.20	
		Spring 2014	29.50	19.40	0.24	*	29.40	19.30	0.24	*
	Institution									
		1	6.50	10.30	-0.14		7.00	10.40	-0.12	
		2	13.00	17.80	-0.14		13.80	17.70	-0.11	
		3	9.20	21.20	-0.34	*	8.70	20.90	-0.35	*
		4	†	†	†		†	†	†	
		5	†	†	†		†	†	†	
		6	†	†	†		†	†	†	
		7	23.60	26.50	-0.07		23.40	27.10	-0.08	
		8	11.60	8.00	0.12		11.00	7.80	0.11	
		9	11.50	5.90	0.20	*	11.20	6.00	0.19	
		10	†	†	†		†	†	†	
		11	9.90	1.90	0.34	*	9.70	1.90	0.34	*
		12	0.20	0.90	-0.10		0.20	0.90	-0.09	
		13	2.20	1.20	0.08		2.40	1.20	0.09	
		14	12.20	6.20	0.21	*	12.60	6.10	0.22	*

* Indicates the standardized bias is greater than 20%
† Indicates the field is not applicable.

Table 3c. Pre-matching descriptive statistics and standardized differences for relevant covariates by English sample.

			Pre-matching Separate English - Withdraw				Pre-matching Separate English - Grade			
Level	Variable	Value	Treated Mean	Untreated Mean	Standardized Difference		Treated Mean	Untreated Mean	Standardized Difference	
Student by semester by institution	Propensity score		0.40	0.07	1.47	*	0.40	0.07	1.50	*
	Number of credit hours		13.15	14.99	-0.59	*	13.03	14.98	-0.62	*
	Received Pell Grant									
		No	7.90	10.20	-0.08		8.20	10.30	-0.07	
		Yes	12.30	14.00	-0.05		11.90	13.90	-0.06	
		Missing	79.80	75.80	0.10		80.00	75.70	0.10	
	Academic level (undergraduate)									
		First year	7.20	53.60	-1.17	*	6.90	53.70	-1.18	*
		Second year	18.40	23.90	-0.13		17.90	23.90	-0.15	
		Third year	32.20	13.40	0.46	*	32.20	13.20	0.47	*
		Fourth year	38.90	8.00	0.78	*	39.40	8.00	0.80	*
		Fifth year	0.20	0.00	0.04		0.20	0.00	0.05	
		Unknown	3.10	1.20	0.14		3.40	1.20	0.15	
Student by semester	Permanent Residence									
		Urban	51.40	52.10	-0.02		51.30	52.10	-0.02	
		Rural	42.20	34.70	0.16		42.20	34.60	0.16	
		Unknown	6.40	13.20	-0.23	*	6.50	13.30	-0.23	*
	Dependency status									
		Dependent	10.90	22.60	-0.32	*	10.60	22.70	-0.33	*
		Independent	9.30	1.60	0.35	*	9.40	1.60	0.35	*
		Missing	79.80	75.80	0.10		80.00	75.70	0.10	
	Income percentile									
		First percentile	5.80	4.50	0.06		5.60	4.50	0.05	
		Second percentile	5.40	5.10	0.02		5.30	5.10	0.01	
		Third percentile	3.50	5.20	-0.08		3.60	5.20	-0.08	
		Fourth percentile	5.00	9.30	-0.17		5.10	9.30	-0.17	
		Missing	80.20	75.90	0.11		80.40	75.90	0.11	
	UNC GPA percentile									
		First percentile	6.70	3.90	0.13		6.60	3.70	0.13	
		Second percentile	5.00	3.30	0.09		4.80	3.30	0.07	
		Third percentile	4.90	3.20	0.09		4.90	3.20	0.09	
		Fourth percentile	4.70	2.50	0.12		5.00	2.50	0.13	
		Missing	78.70	87.20	-0.23	*	78.70	87.30	-0.23	*

			Pre-matching Separate English - Withdraw				Pre-matching Separate English - Grade			
Level	Variable	Value	Treated Mean	Untreated Mean	Standardized Difference		Treated Mean	Untreated Mean	Standardized Difference	
Student	Age		25.19	19.00	0.92	*	25.28	19.00	0.93	*
	Gender	Male	0.32	0.40	-0.19		0.32	0.40	-0.17	
	High school GPA percentile									
		First percentile	19.10	24.80	-0.14		18.20	24.50	-0.16	
		Second percentile	15.50	26.30	-0.27	*	15.40	26.30	-0.27	*
		Third percentile	15.20	22.40	-0.19		15.40	22.60	-0.18	
		Fourth percentile	13.00	16.50	-0.10		13.40	16.70	-0.09	
		Missing	37.30	10.00	0.68	*	37.70	9.90	0.69	*
	SAT score percentile									
		First percentile	12.50	22.30	-0.26	*	11.80	22.40	-0.29	*
		Second percentile	11.30	22.70	-0.31	*	11.20	22.70	-0.31	*
		Third percentile	12.70	18.40	-0.16		12.70	18.30	-0.16	
		Fourth percentile	12.80	16.40	-0.10		12.90	16.40	-0.10	
		Missing	50.70	20.30	0.67	*	51.40	20.20	0.69	*
Enrollment (student by semester by institution)	Semester									
		Fall 2012	26.50	26.70	-0.01		26.50	26.90	-0.01	
		Spring 2013	25.80	24.60	0.03		26.50	24.50	0.05	
		Fall 2013	23.20	25.60	-0.06		22.90	25.80	-0.07	
		Spring 2014	24.50	23.10	0.03		24.00	22.90	0.03	
	Institution									
		1	7.50	4.10	0.15		7.80	4.10	0.15	
		2	25.40	36.50	-0.24	*	26.20	36.60	-0.23	*
		3	†	†	†		†	†	†	
		4	5.10	3.70	0.07		5.20	3.60	0.08	
		5	1.80	0.40	0.13		1.60	0.40	0.12	
		6	4.60	0.40	0.27	*	4.80	0.40	0.28	*
		7	16.70	5.20	0.37	*	17.00	5.10	0.39	*
		8	9.60	14.90	-0.16		7.80	14.90	-0.22	*
		9	8.70	5.00	0.14		8.40	4.90	0.14	
		10	5.30	9.20	-0.15		5.40	9.40	-0.15	
		11	8.70	4.70	0.16		8.90	4.70	0.17	
		12	2.40	5.10	-0.14		2.50	5.10	-0.14	
		13	0.50	0.10	0.08		0.50	0.10	0.08	
		14	3.80	10.70	-0.27	*	3.90	10.60	-0.26	*

* Indicates the standardized bias is greater than 20%
† Indicates the field is not applicable.

Table 4a. Post-matching descriptive statistics and standardized differences for relevant covariates by overall sample.

Level	Variable	Value	Post-matching Overall - Withdraw Treated Mean	Post-matching Overall - Withdraw Untreated Mean	Post-matching Overall - Withdraw Standardized Difference		Post-matching Overall - Grade Treated Mean	Post-matching Overall - Grade Untreated Mean	Post-matching Overall - Grade Standardized Difference	
Student by semester by institution										
	Propensity score		0.32	0.24	0.33	*	0.32	0.24	0.32	*
	Number of credit hours		13.24	14.00	-0.20		13.15	13.97	-0.21	*
	Received Pell Grant									
		No	7.70	7.00	0.02		7.90	7.20	0.03	
		Yes	12.10	12.60	-0.02		11.90	12.40	-0.02	
		Missing	80.20	80.30	0.00		80.20	80.40	-0.01	
	Academic level (undergraduate)									
		First year	10.30	13.30	-0.09		10.00	12.80	-0.09	
		Second year	22.40	26.40	-0.09		22.10	26.00	-0.09	
		Third year	30.80	32.60	-0.04		30.60	32.50	-0.04	
		Fourth year	32.70	25.50	0.16		33.20	26.40	0.15	
		Fifth year	0.40	0.40	0.00		0.30	0.30	0.00	
		Unknown	3.50	1.90	0.10		3.80	2.00	0.11	
	Course subject									
		English	55.20	52.20	0.06		55.90	52.80	0.06	
		Math	44.80	47.80	-0.06		44.10	47.20	-0.06	
Student by semester										
	Permanent Residence									
		Urban	54.10	56.10	-0.04		53.80	55.90	-0.04	
		Rural	39.00	37.70	0.03		39.20	37.90	0.03	
		Unknown	6.90	6.20	0.03		7.00	6.20	0.03	
	Dependency status									
		Dependent	11.40	13.70	-0.07		11.30	13.60	-0.07	
		Independent	8.40	6.00	0.09		8.50	6.00	0.10	
		Missing	80.20	80.30	0.00		80.20	80.40	-0.01	
	Income percentile									
		First percentile	5.50	5.30	0.01		5.50	5.10	0.02	
		Second percentile	5.10	5.00	0.01		5.10	5.00	0.01	
		Third percentile	3.70	3.80	-0.01		3.60	3.90	-0.01	
		Fourth percentile	4.90	5.30	-0.02		5.00	5.30	-0.01	
		Missing	80.70	80.60	0.00		80.70	80.70	0.00	
	UNC GPA percentile									
		First percentile	5.80	6.30	-0.02		5.70	6.20	-0.02	
		Second percentile	5.00	5.70	-0.03		4.70	5.60	-0.04	
		Third percentile	4.40	4.70	-0.01		4.40	4.80	-0.02	
		Fourth percentile	3.80	2.90	0.05		3.90	3.00	0.05	
		Missing	81.00	80.40	0.02		81.20	80.40	0.02	

			Post-matching Overall - Withdraw				Post-matching Overall - Grade			
Level	Variable	Value	Treated Mean	Untreated Mean	Standardized Difference		Treated Mean	Untreated Mean	Standardized Difference	
Student										
	Age		24.83	22.89	0.22	*	24.88	22.92	0.22	*
	Gender	Male	0.34	0.36	-0.06		0.33	0.36	-0.05	
	High school GPA percentile									
		First percentile	17.60	17.60	0.00		16.90	16.80	0.00	
		Second percentile	15.10	17.60	-0.07		15.20	17.80	-0.07	
		Third percentile	15.30	19.10	-0.10		15.40	19.00	-0.10	
		Fourth percentile	12.60	14.30	-0.05		12.90	14.60	-0.05	
		Missing	39.40	31.40	0.17		39.70	31.70	0.17	
	SAT score percentile									
		First percentile	12.90	15.50	-0.07		12.60	14.90	-0.07	
		Second percentile	11.60	14.60	-0.09		11.60	14.60	-0.09	
		Third percentile	12.20	14.60	-0.07		12.20	14.80	-0.08	
		Fourth percentile	12.40	14.10	-0.05		12.40	13.90	-0.04	
		Missing	50.80	41.30	0.19		51.20	41.90	0.19	
Enrollment (student by semester by institution)										
	Semester									
		Fall 2012	24.50	23.20	0.03		24.60	23.10	0.04	
		Spring 2013	27.00	28.00	-0.02		27.60	28.50	-0.02	
		Fall 2013	22.00	22.30	-0.01		21.80	22.20	-0.01	
		Spring 2014	26.40	26.50	0.00		26.00	26.20	0.00	
	Institution									
		1	7.30	7.30	0.00		7.70	7.70	0.00	
		2	18.70	18.70	0.00		19.60	19.60	0.00	
		3	4.20	4.20	0.00		3.90	3.90	0.00	
		4	3.00	3.00	0.00		3.10	3.10	0.00	
		5	1.00	1.00	0.00		0.80	0.80	0.00	
		6	2.70	2.70	0.00		2.90	2.90	0.00	
		7	20.50	20.50	0.00		20.70	20.70	0.00	
		8	10.90	10.90	0.00		9.60	9.60	0.00	
		9	10.10	10.10	0.00		9.90	9.90	0.00	
		10	3.00	3.00	0.00		3.20	3.20	0.00	
		11	9.20	9.20	0.00		9.20	9.20	0.00	
		12	1.20	1.20	0.00		1.20	1.20	0.00	
		13	0.50	0.50	0.00		0.40	0.40	0.00	
		14	7.70	7.70	0.00		7.80	7.80	0.00	

* Indicates the standardized bias is greater than 20%

Table 4b. Post-matching descriptive statistics and standardized differences for relevant covariates by math sample.

			Post-matching Separate Math - Withdraw				Post-matching Separate Math - Grade			
Level	Variable	Value	Treated Mean	Untreated Mean	Standardized Difference		Treated Mean	Untreated Mean	Standardized Difference	
Student by semester by institution										
	Propensity score		0.29	0.20	0.34	*	0.29	0.20	0.33	*
	Number of credit hours		12.86	13.74	-0.21	*	12.78	13.66	-0.21	*
	Received Pell Grant									
		No	6.90	6.80	0.00		7.00	7.00	0.00	
		Yes	11.80	12.90	-0.03		11.60	13.10	-0.04	
		Missing	81.30	80.30	0.03		81.40	79.90	0.04	
	Academic level (undergraduate)									
		First year	13.90	18.00	-0.11		13.70	18.00	-0.12	
		Second year	26.50	30.20	-0.08		26.60	30.10	-0.08	
		Third year	28.70	28.80	0.00		28.30	28.80	-0.01	
		Fourth year	26.00	20.10	0.14		26.20	20.00	0.15	
		Fifth year	0.60	0.40	0.03		0.50	0.30	0.03	
		Unknown	4.20	2.50	0.09		4.60	2.70	0.10	
Student by semester										
	Permanent Residence									
		Urban	57.30	60.40	-0.06		56.90	60.30	-0.07	
		Rural	35.60	34.00	0.03		35.80	33.90	0.04	
		Unknown	7.10	5.60	0.06		7.30	5.80	0.06	
	Dependency status									
		Dependent	11.00	13.30	-0.07		11.20	13.90	-0.08	
		Independent	7.70	6.30	0.05		7.50	6.20	0.05	
		Missing	81.30	80.30	0.03		81.40	79.90	0.04	
	Income percentile									
		First percentile	4.90	5.50	-0.03		4.90	5.30	-0.02	
		Second percentile	4.80	4.90	-0.01		4.60	5.00	-0.02	
		Third percentile	3.90	4.20	-0.01		3.90	4.50	-0.03	
		Fourth percentile	4.50	4.70	-0.01		4.60	4.90	-0.01	
		Missing	81.90	80.70	0.03		81.90	80.20	0.04	
	UNC GPA percentile									
		First percentile	4.40	6.00	-0.07		4.40	5.70	-0.06	
		Second percentile	4.60	5.60	-0.04		4.30	5.40	-0.05	
		Third percentile	3.70	4.50	-0.04		3.60	4.80	-0.06	
		Fourth percentile	2.60	2.00	0.04		2.70	2.10	0.04	
		Missing	84.60	81.80	0.07		85.00	82.00	0.08	

Level	Variable	Value	Post-matching Separate Math - Withdraw Treated Mean	Post-matching Separate Math - Withdraw Untreated Mean	Post-matching Separate Math - Withdraw Standardized Difference		Post-matching Separate Math - Grade Treated Mean	Post-matching Separate Math - Grade Untreated Mean	Post-matching Separate Math - Grade Standardized Difference	
Student										
	Age		25.53	23.33	0.24	*	25.46	23.24	0.25	*
	Gender	Male	0.35	0.37	-0.05		0.35	0.37	-0.05	
	High school GPA percentile									
		First percentile	15.80	16.00	-0.01		15.40	14.90	0.01	
		Second percentile	13.30	16.40	-0.09		13.70	16.20	-0.07	
		Third percentile	14.10	18.70	-0.13		14.00	19.20	-0.14	
		Fourth percentile	10.60	12.80	-0.07		10.80	13.30	-0.08	
		Missing	46.20	36.20	0.21	*	46.20	36.40	0.20	
	SAT score percentile									
		First percentile	13.10	16.00	-0.08		13.10	15.80	-0.08	
		Second percentile	11.00	13.50	-0.08		11.00	13.60	-0.08	
		Third percentile	10.50	12.80	-0.07		10.40	12.40	-0.06	
		Fourth percentile	10.70	12.50	-0.06		10.70	12.60	-0.06	
		Missing	54.70	45.20	0.19		54.80	45.50	0.19	
Enrollment (student by semester by institution)										
	Semester									
		Fall 2012	21.50	21.40	0.00		21.30	21.30	0.00	
		Spring 2013	28.70	29.40	-0.02		29.10	29.60	-0.01	
		Fall 2013	20.70	19.10	0.04		20.50	19.20	0.03	
		Spring 2014	29.20	30.10	-0.02		29.00	29.90	-0.02	
	Institution									
		1	6.70	6.70	0.00		7.10	7.10	0.00	
		2	13.00	13.00	0.00		13.90	13.90	0.00	
		3	9.40	9.40	0.00		8.90	8.90	0.00	
		4	†	†	†		†	†	†	
		5	†	†	†		†	†	†	
		6	†	†	†		†	†	†	
		7	24.00	24.00	0.00		23.80	23.80	0.00	
		8	11.70	11.70	0.00		11.10	11.10	0.00	
		9	11.70	11.70	0.00		11.40	11.40	0.00	
		10	†	†	†		†	†	†	
		11	10.10	10.10	0.00		9.90	9.90	0.00	
		12	0.20	0.20	0.00		0.20	0.20	0.00	
		13	1.00	1.00	0.00		1.10	1.10	0.00	
		14	12.30	12.30	0.00		12.80	12.80	0.00	

* Indicates the standardized bias is greater than 20%
† Indicates the field is not applicable.

Table 4c. Post-matching descriptive statistics and standardized differences for relevant covariates by English sample.

			Post-matching Separate English - Withdraw				Post-matching Separate English - Grade			
Level	Variable	Value	Treated Mean	Untreated Mean	Standardized Difference		Treated Mean	Untreated Mean	Standardized Difference	
Student by semester by institution										
	Propensity score		0.37	0.27	0.35	*	0.38	0.28	0.35	*
	Number of credit hours		13.49	14.28	-0.22	*	13.39	14.26	-0.24	*
	Received Pell Grant									
		No	8.20	6.90	0.05		8.40	6.90	0.06	
		Yes	12.60	12.60	0.00		12.20	12.10	0.00	
		Missing	79.20	80.60	-0.03		79.40	81.00	-0.04	
	Academic level (undergraduate)									
		First year	7.50	8.70	-0.04		7.20	7.90	-0.03	
		Second year	19.00	23.20	-0.10		18.60	22.70	-0.10	
		Third year	32.10	36.50	-0.09		32.20	36.80	-0.10	
		Fourth year	38.40	30.20	0.17		39.00	31.20	0.16	
		Fifth year	0.20	0.20	-0.01		0.20	0.20	-0.01	
		Unknown	2.70	1.10	0.11		2.90	1.20	0.12	
Student by semester										
	Permanent Residence									
		Urban	51.70	53.40	-0.03		51.60	53.70	-0.04	
		Rural	41.70	39.50	0.04		41.80	39.10	0.05	
		Unknown	6.60	7.00	-0.02		6.60	7.20	-0.02	
	Dependency status									
		Dependent	11.40	13.00	-0.05		11.20	12.70	-0.05	
		Independent	9.40	6.40	0.11		9.50	6.30	0.12	
		Missing	79.20	80.60	-0.03		79.40	81.00	-0.04	
	Income percentile									
		First percentile	6.10	6.00	0.00		5.90	5.40	0.02	
		Second percentile	5.60	4.80	0.04		5.40	5.10	0.02	
		Third percentile	3.50	3.10	0.02		3.50	3.00	0.03	
		Fourth percentile	5.20	5.40	-0.01		5.20	5.40	-0.01	
		Missing	79.70	80.70	-0.03		79.90	81.10	-0.03	
	UNC GPA percentile									
		First percentile	6.70	6.40	0.01		6.60	6.30	0.01	
		Second percentile	5.30	6.30	-0.05		5.00	6.10	-0.05	
		Third percentile	5.00	5.30	-0.01		5.00	5.40	-0.02	
		Fourth percentile	4.80	3.40	0.07		5.10	3.70	0.07	
		Missing	78.20	78.50	-0.01		78.20	78.50	-0.01	

Level	Variable	Value	Post-matching Separate English - Withdraw				Post-matching Separate English - Grade			
			Treated Mean	Untreated Mean	Standardized Difference		Treated Mean	Untreated Mean	Standardized Difference	
Student										
	Age		24.75	22.56	0.26	*	24.84	22.65	0.25	*
	Gender	Male	0.32	0.34	-0.05		0.32	0.33	-0.04	
	High school GPA percentile									
		First percentile	19.10	20.40	-0.03		18.30	18.90	-0.02	
		Second percentile	15.90	18.20	-0.06		15.80	18.20	-0.07	
		Third percentile	15.60	18.60	-0.08		15.80	19.30	-0.09	
		Fourth percentile	13.40	15.00	-0.05		13.80	15.40	-0.05	
		Missing	36.00	27.80	0.18		36.40	28.20	0.18	
	SAT score percentile									
		First percentile	12.60	16.20	-0.10		12.00	15.00	-0.09	
		Second percentile	11.60	14.00	-0.07		11.50	14.00	-0.08	
		Third percentile	13.10	15.60	-0.07		13.10	16.00	-0.08	
		Fourth percentile	13.10	15.10	-0.06		13.10	15.50	-0.07	
		Missing	49.70	39.10	0.21	*	50.30	39.50	0.22	*
Enrollment (student by semester by institution)										
	Semester									
		Fall 2012	26.80	25.90	0.02		27.00	25.50	0.03	
		Spring 2013	25.60	26.20	-0.01		26.30	26.90	-0.01	
		Fall 2013	23.10	25.20	-0.05		22.80	25.10	-0.06	
		Spring 2014	24.40	22.70	0.04		24.00	22.50	0.04	
	Institution									
		1	7.80	7.80	0.00		8.10	8.10	0.00	
		2	23.80	23.80	0.00		24.50	24.50	0.00	
		3	†	†	†		†	†	†	
		4	5.30	5.30	0.00		5.40	5.40	0.00	
		5	1.80	1.80	0.00		1.60	1.60	0.00	
		6	4.90	4.90	0.00		5.10	5.10	0.00	
		7	17.50	17.50	0.00		17.90	17.90	0.00	
		8	9.90	9.90	0.00		8.10	8.10	0.00	
		9	8.90	8.90	0.00		8.70	8.70	0.00	
		10	5.40	5.40	0.00		5.50	5.50	0.00	
		11	8.70	8.70	0.00		8.90	8.90	0.00	
		12	1.90	1.90	0.00		2.00	2.00	0.00	
		13	0.10	0.10	0.00		0.20	0.20	0.00	
		14	3.90	3.90	0.00		3.90	3.90	0.00	

* Indicates the standardized bias is greater than 20%
† Indicates the field is not applicable.

Table 5. Average treatment effects on the treated by outcome model and sample.

	All enrollments	Math enrollments	English enrollments
Outcome: Withdraw			
Effect estimate	0.0573***	0.0365***	0.0591***
	(0.0048)	(0.0069)	(0.0052)
Observations	14,839	7,923	9,754
Outcome: Grade			
Effect estimate	-0.1813***	-0.1110***	-0.2752***
	(0.0215)	(0.0343)	(0.0263)
Observations	13,227	6,087	7,392

Standard errors in parentheses
*** p<0.01

Figure 1. Rovai's (2003) model of online student persistence.

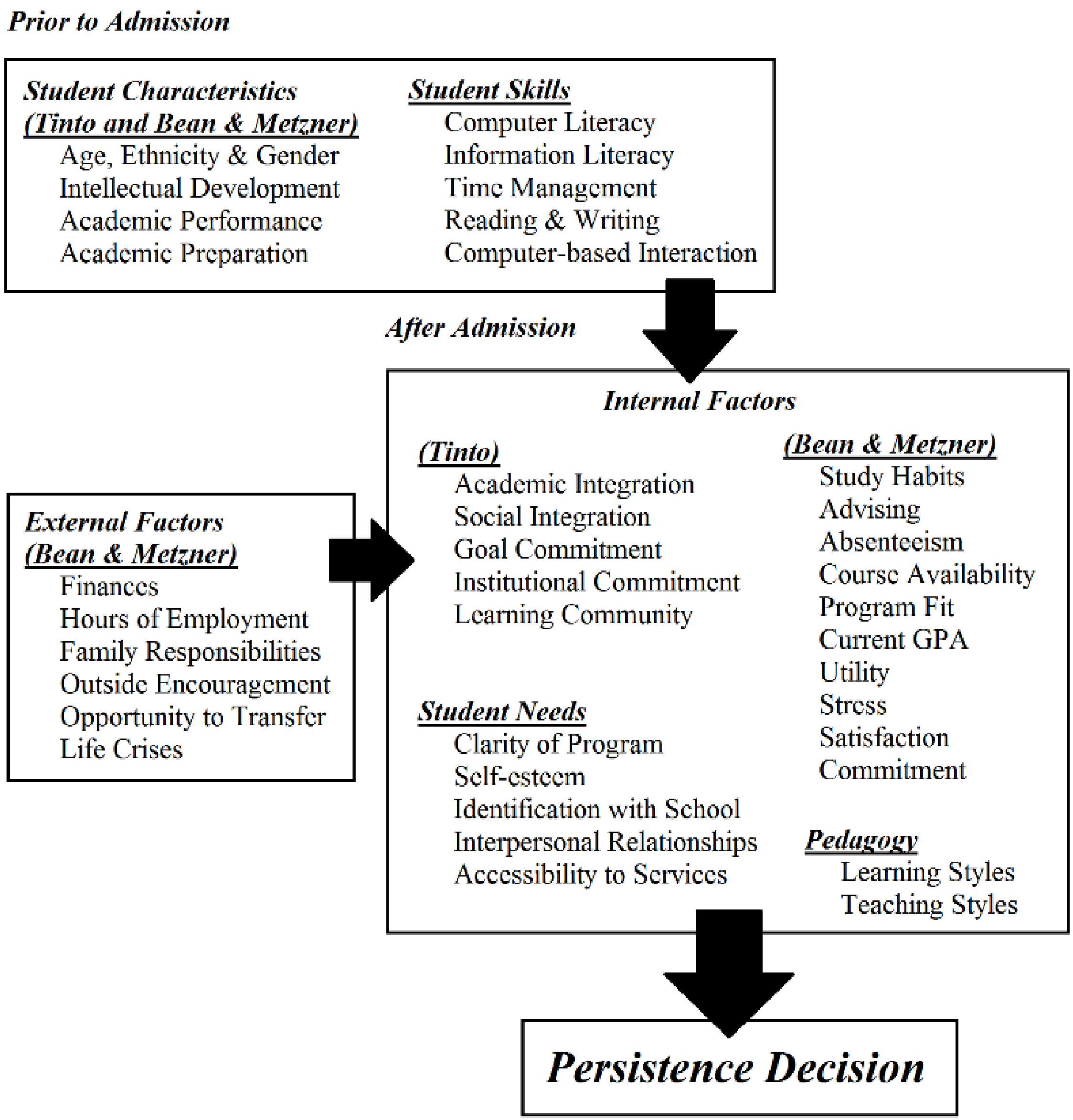

Figure 2. Example of Mahalanobis distance edit for missing values.

Scenario	ID	Course	Online mode of delivery	SAT Score (normalized)	SAT missing indicator	Mahalanobis distance
A	1	ENG101	0	0.45	0	0.45
	2	ENG101	1	0	0	
B	1	ENG101	0	0.45	0	1
	2	ENG101	1	·	1	
C	1	ENG101	0	·	1	0
	2	ENG101	1	·	1	

Figure 3a. Difference in pre-matching standardized bias and post-matching standardized bias for select covariates for the full sample of overall enrollments.

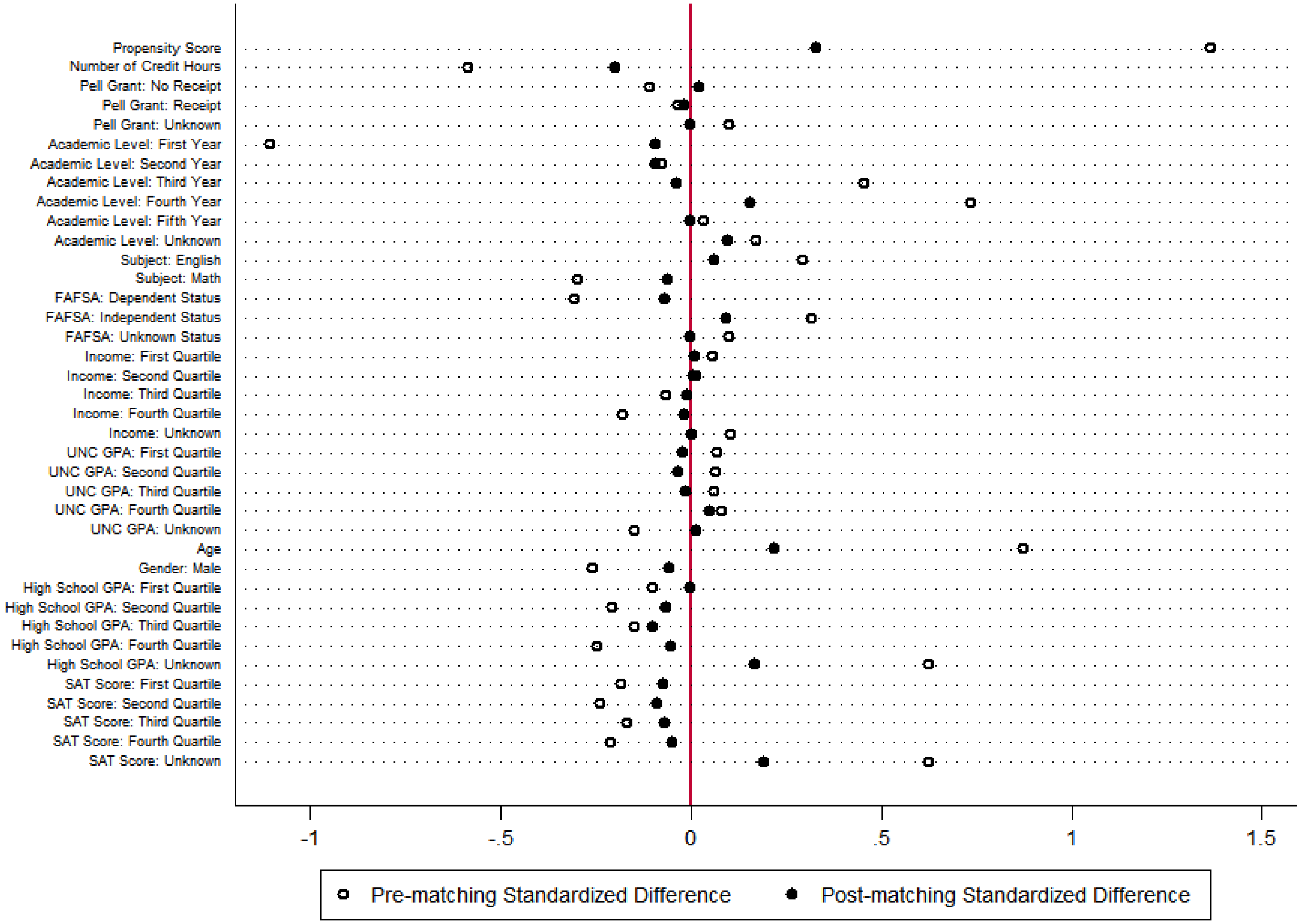

Figure 3b. Difference in pre-matching standardized bias and post-matching standardized bias for select covariates for the sample of math enrollments.

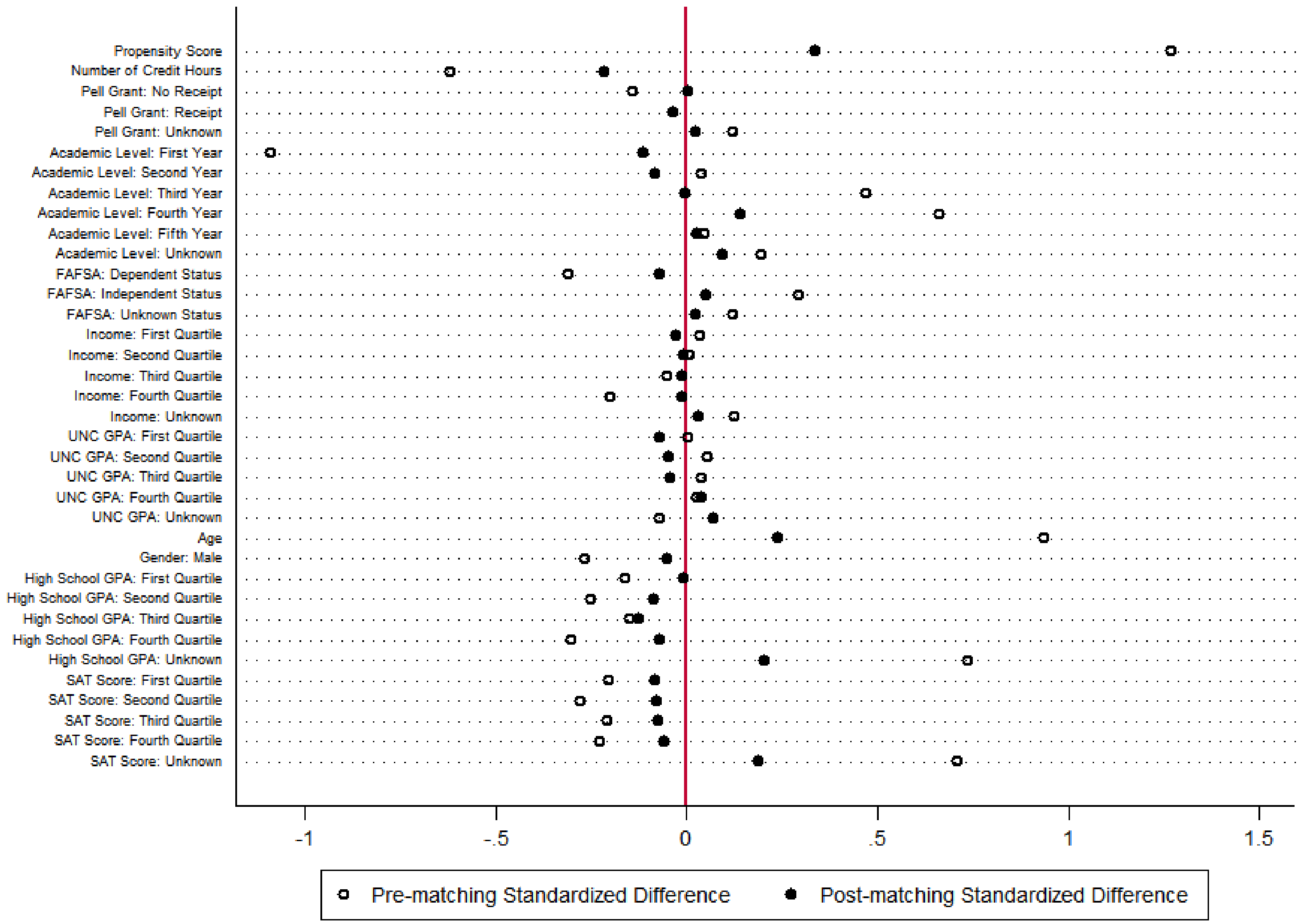

Figure 3c. Difference in pre-matching standardized bias and post-matching standardized bias for select covariates for the sample of English enrollments.