Successful Online Courses in California's Community Colleges

Technical Appendices

CONTENTS

Hans Johnson, Marisol Cuellar Mejia, and Kevin Cook

June 2015

Appendix A: Data and Methods

Data and definitions

The data used in this report come from the California Community College System Chancellor's Office Management Information System (COMIS). Specifically, we use a longitudinal data set of students enrolled in each one of the 112 colleges in the California Community College System. This data set includes student demographics, course enrollment, transcript data, and information on each course (including subject, credit, transfer, and basic skills status) and covers the period from Fall 2006 through Spring 2014.[1]

We categorize community college courses into three groups:

- Online courses are those in which at least 80 percent of the instruction is Internet-based. These courses are accessed online through the Internet, and are classified either as "delayed interaction" courses, in which the student accesses the course at times that are convenient for the student, or as "simultaneous interaction" courses, in which the student must be online at the same time as the instructor. The vast majority of online courses (91% in 2011–12) are offered as delayed interaction. About 11 percent of credit course enrollment in 2011–12 was in online courses.

- Blended courses are those in which a substantial share of instruction—30 to 80 percent—is offered online, with the remainder of instruction occurring in a traditional classroom setting (sometimes referred to as "face-to-face"). The manner in which courses are blended varies: some courses use online instruction for lectures and traditional settings for labs, and others offer a mix of online and traditional settings for lectures and sections. Only 1 percent of credit enrollment in 2011–12 was in blended courses. Because blended courses are so uncommon, we do not examine them in any detail in this report.

- Traditional courses are those in which the vast majority of instruction (70% or more) takes place in a traditional classroom with the instructor and students physically present at the same time. Technology is often used in such classrooms, including Internet and online components, but primarily within the classroom itself. In 2011–12, 87 percent of total credit enrollment was in traditional courses.

One challenge of the data is that colleges use different course numbers and titles for the same course (e.g., Introduction to Managerial Accounting will be Accounting 110 at one college and Accounting 100 at another college). To identify common courses across colleges, we use the "Course Identification Numbering System" (CID). CID is a supranumbering system being developed to ease the transfer and articulation burdens in California's higher educational institutions. This project is funded by the California Community College System Office and is a collaborative effort of the Academic Senate of the California Community Colleges, the Academic Senate of the California State University, the Academic Senate of the University of California, and the Association of Independent California Colleges and Universities. CID addresses the need for "common course numbers" by providing a mechanism to identify comparable courses. Most CID numbers identify lower-division transferable courses commonly articulated between the California community colleges and universities (including Universities of California, the California State Universities, as well as with many of California's independent colleges and universities). The CID number is a designation that ties that course to a specific course "descriptor" that was developed by intersegmental discipline faculty and reviewed statewide. CID began by developing descriptors for courses in 20 disciplines that are among those most frequently transferred. The next area of focus has been on the courses in TMC's (the Transfer Model Curricula that describe the major component of associate degrees for transfer). CID will then expand to include more and more courses each year.

---

[1] The COMIS was implemented in 1990 and seeks to collect data that can provide answers to fundamental questions related to students, courses, outcomes, student services, and faculty and staff. Colleges submit data to the Chancellor's Office within 30 days of the end of each term.

Unfortunately, a linking field was not developed between the CID database and the COMIS database to cross-link these two together. COMIS does not collect CID numbers. CID does not collect unique COMIS course numbers. Alpha fields are used for course identification. So, ENG101 in COMIS can be ENGL 101 in CID. To overcome this challenge, we create a linking filed using a multi-step approach. We are able to identify in the COMIS database over 90% of the college-specific courses with a CID assigned. We use information on CID's as of September 6th 2014, which includes 31 disciplines, 220 descriptors, and 9,143 courses.[2]

For our statistical models, we restrict our sample to include:

- Courses offered between 2009–2010 and 2013–14 academic years.
- Courses with a valid CID.
- Courses offered through both online and face-to-face sections.

After applying all the above restrictions, our final sample consisted of 1,678 courses. Each one of these courses is offered through multiple sections (on average six face-to-face sections and two online sections in a single term). Our final database includes 48,891 face-to-face sections and 16,541 online sections. Student enrollment in these courses exceeds 1.6 million.

Methods

Identifying high-quality online courses poses multiple challenges. First, as noted above, there is not a common course numbering system across community colleges. Second, it is possible that student success rates are lower in some online courses because those courses attract students who are less academically able. Third, grades are the best measure we have of student learning, but grades are subjective. Some instructors might be relatively easy graders and some courses might be relatively easy in terms of content. Therefore, identifying online courses with strong student outcomes requires controlling for a plethora of factors.

To overcome the limitation imposed by the fact that there is not a common course system, we focus only on courses part of the CID project. To predict the probability of student success in specific course sections controlling for student mix and other factors, we run student-level linear probability models.[3] We include a rich set of controls for student demographic characteristics, including measures of academic skill and preparedness. Specifically, we include dummy variables for college age, gender, race/ethnicity, highest level of education at the time of initial enrollment, prior dual enrollment, prior experience in basic skill courses, full-time status, disability status, academically disadvantaged status, and intent to transfer. We also include the students' GPA in traditional courses. To address the potential problem that online courses may be more prevalent within some colleges, terms, and subject areas, we add fixed effects for term, college of attendance and CID. Standard errors are clustered at the CID level.

We use the results from our models to develop predicted success rates by course section (equal to the sum of the predicted individual student probabilities of success aggregated over all the students in a particular course section). These predicted course-section success rates are developed for both online and traditional settings. Our pool of successful online courses is drawn from the set of online course sections in which the observed course section success rate exceeds the predicted success rate, and where the observed online course section success rate is higher than the observed and predicted success rate in the traditional version of the course.

---

[2] More information is available at www.c-id.net/.

[3] We are using these models for predictive comparisons between groups not to find causal relationships. Therefore selection bias is less of an issue in this context. Our previous report (Johnson and Cuellar-Mejia 2014) was all about finding causal relationships, therefore dealing with self-selection biases was a big part of our empirical strategy.

Finally, to overcome the problem of relatively easy graders, we consider subsequent academic performance of students in other courses in the same subject. If students do very well in the initial course, but poorly in subsequent traditional or online courses, we eliminate the initial course from our list of highly successful courses. Specifically, in the case of courses of the same sequence, the indicator of course quality that we use is the percentage of students who take an online prerequisite and then enroll in the subsequent course and receive a grade at least as high in the subsequent course. For this analysis we restrict the sample to:

- Graded courses.
- Students who took the initial course for the first time (that is, we are eliminating students who are retaking a course for a second time because they failed the first time around).
- Students who completed the course.
- Students who took at least two courses in the same discipline.
- We eliminate those students taking different sections of the same course and achieving a passing grade in both (very few students fall in this category).

We combine observations across terms to achieve a larger sample of students. If we were to focus on a single term / college / course, we would end up drawing conclusions from very small sample sizes.

Even with the methods we employ above, we are not able to fully control for course-level variation in instructional quality and support or grading standards. For example, we were unable to include measures of teacher quality given data constraints.

We report our results from the analyses in the main body of the report ("How Many Online Courses Are Successful?"). We also use our results to identify the very highest performing online courses. These courses formed the basis of our sample for qualitative interviews with faculty and community college officials (including distance-education coordinators). The information obtained in the interviews supplemented our literature review of best practices in online learning. The very highest-ranking courses are shown in Table A1 (names of colleges are omitted for confidentiality).

**TABLE A1**
Highest Performing Online Courses, 2013–14

Course ID (CID)	Course ID (college)	Course title
MECON 102	ECON 201	Principles of Microeconomics
PSYC 7	PSY 150	Introduction to Biological Psychology
ECON 2	ECON 202	Principles of Macroeconomics
POSC 1	POLS 110	Introduction to American Government and Politics
ECON 2	ECON 201	Principles of Microeconomics
PSYCH 11	CDEV 100	Child Growth and Development
CISC 310-1757	ITIS 120	Business Information Systems, Computer Information Systems
ECON 1	ECON 201	Principles of Microeconomics
MPSYCH 101	PSY 110	Introductory Psychology
ACCTG 002	ACCT 120	Managerial Accounting
ECON 2	ECON 201	Principles of Microeconomics
ECE 2	ECE 120	Principles & Practices of Teaching Young Children
BUS 82	BUS 110	Introduction to Business
COMM 121	COMM 130	Interpersonal Communication

Each of these courses met the following criteria:

- Offered in Fall and Spring 2013–14
- Had a traditional course counterpart in each of those semesters.
- Enrollment exceeded 25 students in each semester and in both the online and traditional versions.
- Observed online success rate $> = .70$ (in fall and spring each).
- Observed online success rate $> =$ observed traditional success rate (in fall and spring each).
- Observed online success rate $> =$ predicted online success rate (in fall and spring each). The predicted online success rate is based on our linear probability model.
- Observed performance gap is greater than the predicted performance gap by .15 or more (in fall and spring each). The performance gap is defined as the difference between online performance gap and traditional performance gap; this criteria identifies courses in which online learning success rates were much higher than we expected in comparison with traditional courses.
- Students perform better in subsequent traditional courses than students whose initial counterpart course was a traditional face-to-face course.

Traditional courses tend to have higher passage rates than online courses

Cumulative distributions by passage rate:
online courses and traditional courses
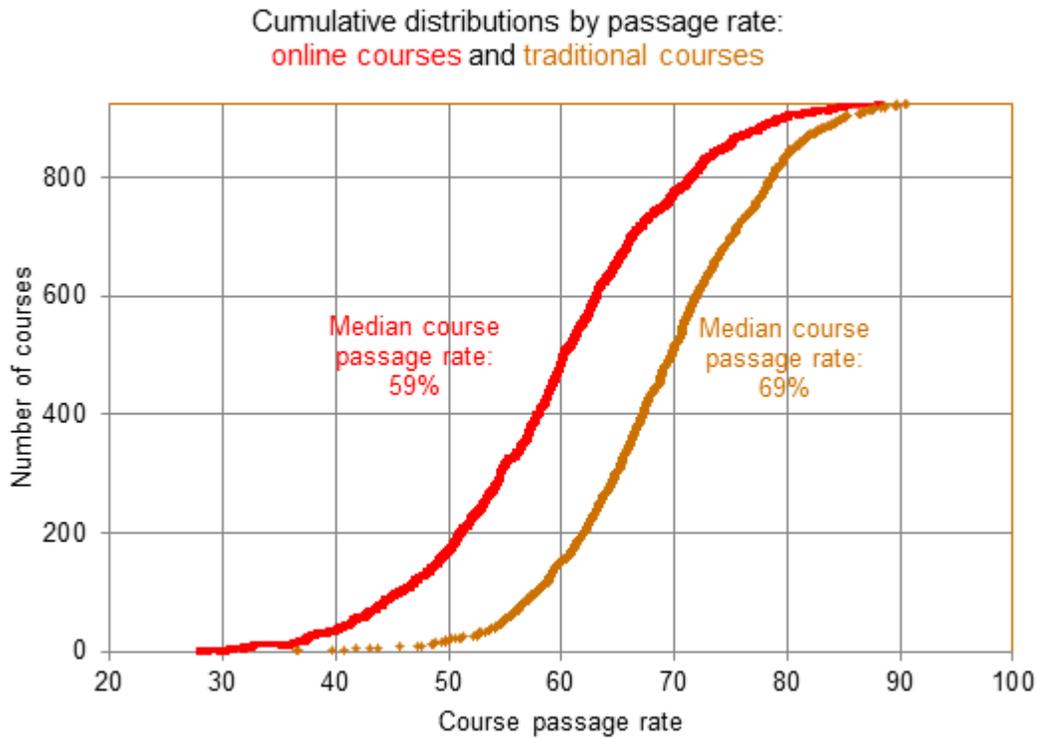
SOURCE: Authors' analysis based on California Community College Chancellor's Office data.

NOTE: Restricted to 2010–2014 courses offered both online and face-to-face in the same college and with at least 250 enrollments in both online and face-to-face sections.

Appendix B: Structured Interview

To inform our understanding of the specific factors that contribute to a successful online course, we conducted eight interviews with online faculty, department chairs, distance-education coordinators, and statewide online education experts. These interviews took place over the course of a few weeks and lasted about one hour each. Given the diversity of expertise and experience of the interviewees, and the exploratory goal of the process, we thought it best to conduct informal phone interviews with open-ended questions that allowed the subjects to share personal experiences and anecdotes to illustrate their answers. On the whole, we were able to identify common themes and shared experiences that helped us better understand our empirical analysis and allowed us to make specific recommendations. Below we present the outline that we used to guide our interviews.

Interview Outline

1. Background
   - Basics
     - Present position, years of experience, field of study
     - How would you describe your role in online teaching, learning, and assessment at this institution?
2. Course design and delivery
   - Briefly describe the online course development process
     - Who decides if a course should be offered online?
     - Who approves an online course?
     - Is training a part of the course development process? Did you receive training? What type of training did you receive?
   - What resources are available for online course design (i.e., support staff, technical help, software programmers, instructional designers)?
   - How have you organized the course content?
     - Is the course content broken into modules? If so, how do you determine the size and breadth of the modules?
     - Do you incorporate various forms of media related to the course content? What type of media do you include—pictures, video, external links?
     - How often do you modify the course content?

3. Learning management system
     - Which learning management system do you use?
     - What would you say are the strengths and weaknesses of your learning management system?
     - If you could make any changes to your school's learning management system what would they be?

4. Screening students
   - Do you screen students before they are allowed to enroll in an online course? If so, what vehicle (test etc.) do you use to screen those students?
   - What share of students pass the screening process?

- Do you provide students with an online orientation before they take an online course?
    - How rigorous is the orientation?
    - How long, on average, does it take to complete the orientation?
  - What do you like about the screening/orientation process? What would you change?

5. Interactivity and Collaboration (Faculty)

- How important is interaction between students and the instructor? Between students and other students?
    - Do you initiate contact with students prior to the beginning of the course?
    - Are student-to-student interactions a course requirement?
    - Do you assess students based on the quality of their interaction; is a rubric provided to students?
    - What type of student-to-student interactions are most/least successful?
        - Synchronous vs. Asynchronous?
        - Do you assign group projects?
- Do you contact students if they seem to be absent or not engaging in the material/interacting with other students?

6. Assessment

  - What type of course material is most successful? What type is least successful?
- How frequently do you assess students' performance in the course?
- How often do you provide students with feedback on their performance?
- What form does this feedback take (grades, written assessment, etc.)?
- If students are struggling with a concept, do you customize the learning modules to suit the individual student?
- Are students given the opportunity to assess the course and instructor?
        - How is this information delivered to the instructor?
        - How often do you modify the course based on student feedback?

7. Student support services

- What support services are available to online students?
        - Is technical support provided? Is it available 24/7?
        - Are institutional services such as a library, writing center, and/or counseling provided to online students? Are links to these services provided within the course format?

8. What did we miss?

- What other types of pedagogical techniques, institutional support, course design, software, or support services make an online course successful?

Have you noticed any trends either positive or negative that have affected the performance of students in the online medium? What are they?

The Public Policy Institute of California is dedicated to informing and improving public policy in California through independent, objective, nonpartisan research.