THE DISTRIBUTION OF LEARNING BENEFITS

FOR WEB-FACILITATED COURSES

BY

HAOTING LUO

A Thesis Submitted to the Faculty of the

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS

In Partial Fulfillment of the Requirements
For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

2011

STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Haoting Luo

Haoting Luo

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

Roger A Dahlgran — 12/2/2011

Dr. Roger Dahlgran — Date

Associate Professor of Agricultural and Resource Economics

ACKNOWLEDGEMENTS

Numerous people have contributed to my success throughout this work, culminating in the production of this thesis. First and foremost, I would like to thank my family for their continuous support over the years.

I have been extremely fortunate to work under the guidance of my amazing advisor, Dr.

Roger Dahlgran, who provided me the opportunity to participate in this exciting project and who was always helping with sincere advice, guidance, support and time on my research, life and career path.

I would like to acknowledge Professor Satheeth Aradhyula and Professor Paul Wilson for their patience and expertise in reading through this work, Nancy Smith, for her continuous assistance and hard work for arranging all the administrative stuff and paperwork for us, and Dr. Rahman Tauhid for his effort and invaluable suggestions on the project. I would also like to thank my classmates who consistently support my study and project.

DEDICATION

To my parents.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Web-facilitated courses are flexible, cost-saving, repeatable and at least as good as traditional education in quality, as discovered by most of the online education researchers. However, defining and optimizing the online learning benefits is still evolving. In this work, we assess the distribution of learning benefits for online teaching by evaluating a combination of recorded data and survey instruments collected from an online course offered at the University of Arizona. We observe a significant positive relationship between the online learning outcome and students' effort in learning the course materials. Factors that positively affect the outcome also include the time spent in traditional classroom learning, students' cumulative academic performance, and other demographic characteristics. In additional, online learning contributes tremendously to examination performance. Determinants of active participation in online learning are also evaluated, and strong correlation with previous learning experiences is observed.

CHAPTER 1 INTRODUCTION

1.1 Introduction

Learning, studying and completing coursework from home, office or other places is no longer impossible. With the dramatic advancement in information technology and the increase in Internet access, schools, universities, and other educational institutions have been putting course-materials online, or delivering entire programs as distance learning opportunities for those who face obstacles to traditional classroom environments.

Online education creates a cyber-environment that provides learners with a resource mix that substitutes for the resources used in a face-to-face environment. Online methodologies offer strengths such as flexibility, savings of some types of costs, duplicability and high quality. Oblinger (2000) identified the advantages of distance education as: 1) increasing the opportunities for learners to access the educational resources; 2) alleviating the cost burdens of constructing physical facilities; 3) discovering new markets and cash flow sources for universities; and 4) providing catalysts for institutional transformation and technological development.

Online learning attracts larger audiences and creates income generating opportunities for schools. Over the past decade, there has been a surge in online learning enrollment’s share of total enrollment at educational institutions. The latest Sloan report on online education (Allen, 2010) reports substantial enrollment in online courses and enrollees having a positive opinion of the future of online education. The number of students taking at least one online course has increased by nearly one million annually and in the fall 2009 term was over 5.6 million students. The annual growth rate in online enrollment of 27% far exceeds the less than 2% annual growth rate of the higher education student population. Nearly thirty percent of all students in higher education have taken an online course. The Sloan research project has been in progress for nearly a decade and reports a collective positive opinion of the future of online education (55.4% positive for private nonprofits and 67.0% positive for for-profits) indicating that online learning is substitutable for and at least as good as face-to-face education in terms of attitudes. The Sloan survey also casts light on the advantage of online education during difficult economic times as online education has experienced a greater demand increase compared to face-to-face offerings.

These findings support the notion that online education has been in the rapid development stage and will play a major role in the future. However, online education is still in its infancy. Issues such as monitoring academic integrity, high drop-rates, and inadequate interaction compared to face-to-face education, and lack of institutional agreements on tuition, quality and evaluation must all be settled. High development costs offset the low maintenance costs, making online education more profitable in the long run but less so in the short run.

With the development of online education, other transitional education methodologies have also been exploited. These include web-facilitated or hybrid methodologies, which are a combination of traditional face-to-face classes and innovative technology supported pedagogies. This suggests that each format has merits and disadvantages.

The Department of Agricultural and Resource Economics at University of Arizona currently offers an undergraduate course entitled "The Economics of Futures Markets".

This course is offered in two formats: as a traditional course using many online tools, and as a developing online course. While revenue generation justifies the development of the online offering, we are interested in the learning benefits attributable to the online tools used in the traditional offering because the online development will rely heavily on these tools. Our interest is the hybrid or web-facilitated course format. The computer-facilitated teaching tools include an online future trading simulation, online homework, online lecture recordings, an online testbank, and online evaluations. Web site interaction data have been recorded in conjunction with each tool developed. These data are valuable for understanding the online learning environment.

Our interest in the following issues will help us build better online educational methods:

Given the present traditional education systems, what kind of distinguishing features can online learning options present to attract more students in this potential market? How do the new-tech education methods supplement learners' efforts? To what extent can learning benefits be achieved purely from the design of online materials?

Given the various learning characteristics of students (e.g. time availability, thinking habits, academic proficiency, clicking behavior), how do different learning styles influence practices and outcomes such as non-participation, deadline rush-hours etc.?

What group-specific characteristics (gender, race, learning style preferences) should be considered in course design to insure that the design is optimized so that online education can be universally beneficial?

1.2 Our Contributions

In this research, we examine students' periodic and final learning outcomes (test scores) and evaluate their relationship to demographic characteristics, learning styles, online studying behaviors, etc. We perform both quantitative and qualitative analysis to assess the distribution of benefits, and calculate the effect of various factors' contributions to the online learning outcomes. By analyzing differences among learners, we delve more deeply into the problem by investigating casual relationships for student's environments and preferences. We also examine how learner's time allocations to each webpage create learning.

The contributions of this work also include technological innovation. It appears to be the first usage-monitoring data collection effort that evaluates students' online learning outcomes based on web access behavior. The data generated are more accurate and complete than the data used in previous research efforts. This work also overcomes the problem of lack of monitoring for online course offerings by using problem templates with randomized answer keys. Under conventional methods, collaborations among students have the potential to embed measurement error in time-score relationships that are critical to evaluating web-based education studies. The data used here do not suffer in this regard.

CHAPTER 2 LITERATURE REVIEW

2.1 Overview

The literature on online education is substantial and increasing. Our search terminology included “online education”, “online learning”, “web-facilitated distance learning”, “web-based teaching”, “e-learning” and “computer-assisted instruction” etc. In the search of relevant literature, the scope covers subjects ranging from learning environments, learners’ outcomes, learners’ characteristics, to administrative aspects (Tallent-Runnels, 2006).

The literature of interest reports on qualitative, quantitative, experiment-based, and non-experiment-based research. Qualitative studies use limited quantitative data to describe the current situation and focus on policy implications and insights. Quantitative research relies on massive data sets and seeks causal relationships between variables. In these studies, researchers design classes in operational environments, observe the outcomes based on recorded data or survey instruments, and perform comparative econometric regressions based on these data.

Researchers have also examined data that range from school environment, student characteristics, to learning outcomes. Environment factors and individual characteristics such as demographic characteristics are mostly collected from school records, and the student’s preferred learning style are from self-reported data. Learning outcomes can be measured by several indicators. Test scores are frequently used. Other measures include detailed scores as dictated by educational theory, and students’ self-reported survey responses measuring their evaluation of the learning process.

The research methodologies in the literature vary according to the data analyzed yet the trend in the literature is from qualitative assessment to quantitative analysis. Most qualitative research illustrates the state of online education at a given point in time. Qualitative methods include descriptive characterization, non-experimental analysis which attributes group-wise or behavioral characteristics to the final results, and experimental analysis where variables are controlled, and results are contrasted owing to differences in experimental conditions.

Brown and Liedholm (2002) evaluated the distribution of learning outcomes for an online undergraduate basic microeconomic course. They discovered significant disadvantages among women and blacks in obtaining the online education. They also report positive correlation between the students' performances and original SAT score and between students' performances and GPA. This research emphasizes the importance of considering the distribution of outcomes. This is a typical non-experimental analysis that describes learning differences among groups. This type of study does not take individual student's learning behavior and preferences into account. These refinements would be helpful in understanding how to improve online educational methodologies.

Gratton-Lavoie and Stanley (2009), Sosin et al (2004), Savage (2009), Ragan (2010), Motiwalla and Tello (2000) use differing datasets and econometric methods to evaluate the effects of demographic characteristics, attendance rates, learners' talent levels, class sizes, and teachers' effects in online learning outcomes.

A vast number of similar studies use similar data and perform similar analyses, but differ based on course subject (Ragan and Walia (2010), Jensen et al (2007), Dutton et al (2001)). The conclusions of these studies are subject matter specific.

Most of these studies show that web-based education performs slightly better on average than traditional face-to-face learning. Online learning combined with traditional methods usually produces better learning outcomes than purely online education. The effect of online education is larger for collaborative instruction than for modes that involve independent study. Generally, positive results have been reported for the introduction of online courses.

2.2 Existing Issues

Despite effort toward improving the quality of cyber-offered courses, several issues remain to be addressed.

These issues are mainly due to experimental scope limitations and a lack of analysis of behavioral variation among learners. Behavioral variation includes course participation, and learning habits. First, most studies are conducted by using the resources at hand, say, the data collected from a certain course or program offered by a particular institution. This kind of experiment does not include discrepancies among regions, subjects, student groups and course policies. As a result, findings are usually applicable to only a certain type of course or situation, rendering it less helpful to other online program initiators. Thus, the benefit of online learning is hard to quantify much less optimize.

Second, even when research covers a variety of courses, these studies fail to consider variation of learners’ behaviors. Although a substantial number of papers account for students’ demographics while evaluating the course delivery process, these differences may be of less use to the improvement of online course design than a tighter focus on learners’ preferred learning styles.

Third, large scale studies ignore the causal relationship between the effort spent in online learning and the overall learning outcome. The current studies fail to distinguish the advantages of online from traditional pedagogies. Without analyzing the direct connection between a pedagogy's use and its learning outcome, any optimization efforts directed toward web-based education might run counter to their intended effects.

Yet, certain studies address the above issues and provide some understanding of the processes at work.

One study focuses on cross sectional comparison between online and face-to-face formats (Sonnenwald, 2003). Although its data are from survey instruments which capture students' learning style preferences and perceptions of technological advantages of different teaching methods, this study finds that that the learning outcomes from web courses are comparable among the many subjects in general education courses. What's more, distinguishing learners by learning styles rather than demographic groups provides a novel and in-depth approach to understanding the relationship between learning activities and outcomes.

Young (2004) reports the results of exploring online pedagogical models based on a "school of all" project by looking at the winners of an online education community contest selected from over 2300 courses. This study highlights the attributes of online courses judged as good across subjects, locations, experiences, and social status. This study is of a qualitative nature so it lacks intensive data analysis. As a result, it does not identify causal relationships between contributing factors and learning outcomes.

Another universal study was conducted by the U.S. Department of Education (USDE) (Means, Toyama, Murphy, Bakia, and Jones, 2009). They performed a meta-analysis for the collective effects of 50 independent web-based learning projects from 1996 to 2008. This study applied econometric analysis to research literature findings to reach overall conclusions based on a broad scope of courses, programs, and targeted student groups. These findings raise researchers' awareness of current trends in online course design and development, and the factors that reflect good educational performance. This novel approach helps evaluate the current situation and serves as a guideline for future initiates.

Dutton, Dutton and Perry (2001) report on the benefits of delivering online undergraduate C++ courses to different student groups. They fit final grades to demographic dummies and an "effort" variable where effort is measured as the final homework score recognizing that homework can be submitted many times before the full score is attained. The process for reaching this homework score is not recorded. They observe significant correlation between the learning outcome and "effort".

Taraban, Maki and Rynearson (1999) measure the distribution of time spent online by students learning a specific concept using log-in time and surfing time for an online learning interface. However, they did not relate the learning outcome with the time distributions, which would have been a better way to evaluate the students' effort in learning.

2.3 This Study's Contribution

This study's contribution is its analysis of user-level behavior and learning pattern data. We also analyze web-facilitated learning as an economic process with inputs, and outputs that can be quantified and analyzed. What's more, we subject our data to econometric analysis so that better and sounder solutions can be proposed.

CHAPTER 3 DATA AND METHODOLOGY

3.1 Background

This work focuses on the learning contributions of online teaching tools used in AREC313: *The Economics of Future Markets* taught in the Department of Agricultural and Resource Economics. Dr. Roger Dahlgran is the sole instructor and has taught the course for over 20 years. Beginning in the early 1990s, AREC313 was developed into a web-facilitated course. Initial efforts were devoted to developing a future market trading simulator. This expertise was then applied to web delivered and graded homework assignments and course evaluations. This effort is an innovation of teaching technology in the Department of Agricultural and Resource Economics.

In conjunction with course development, much useful information about the course environment, materials and learners has been recorded. Since the course-website is self-designed and administered, the course developer has the flexibility and freedom to record whatever information is deemed useful for troubleshooting and monitoring the website's usage. These data are incidentally useful for evaluating the technology's effectiveness. Throughout the deployment of successive innovations, the data collection and document efforts have been maintained.

The data are documented by retained course materials, assignments, tests, student profiles, survey instruments, test scores, online access logs, and credits gained along the course of learning. .

For this research, students' periodic learning outcomes (homework scores, test scores) and behavior records are used as dependent variables. Explanatory data include demographic characteristics, learning style preferences, and online studying behaviors.

Regression analysis is used to analyze variables' contributions to the learning outcomes and accounts for variation among groups and types of behaviors.

3.2 Data Sources

Our data consists of information on test scores, course material browsing history, personal information from university records, website click history, and students' evaluation survey data. The initial dataset contains 146 records of which 84 provided personal information via an evaluative survey. This restricts our access to the complete information for evaluating the web-facilitated course to a certain extent, we looked at the distribution of the class participation and final exam scores for both responders and non-responders' groups. For responders and non-responders, final exam scores averaged 16.59 and 15.94 with standard deviations of 4.25 and 3.96, and quiz point averages of 11.39 and 8.26 with standard deviations of 3.94 and 4.15 respectively. Table 1 summarizes the data availability. Appendix A presents summary statistics and Appendix B shows pie charts of the components of sample observations.

Table 1. Breakdown of Student Records

Total Records	**146**
Administrator and TA	2
Students in online section	12
Students who dropped course	7
Dummy records	13
Students enrolled in target course	**112**
Missing data on personal info survey	28
Final Total	**84**
	75%

Each data record consists of demographic characteristics, online course behavior, scores for each examination item, scores for each part of each online homework assignment, and attendance measured by the number of pop quizzes taken. Student scores come from the online assignments, paper-based multiple choice examinations, and pop quiz scores.

The course website records each student's website interaction. The recorded data consists of log-in statistics, surfing time, intermediate answer-checking interactions, and submit history for the online assignments. Students can test-submit each assignment multiple times so that each student's learning effort is observable.

Besides test scores and online activities, we also have also recorded students' personal information such as their demographic group and employment status. These data come from university records and student surveys. The data include student gender, class (junior, senior, etc.), minority group membership, cumulative grade point average, employment status, time worked per week. We believe that these characteristics will help us better understand students' learning outcomes. The next section describes how those data are generated.

3.2.1 Online Homework Assignments

Two types of online homework are assigned in this course. Problem sets are designed to reinforce concepts, skills, and principles. Futures trading case studies mimic the real-world scenarios relevant to futures trading and are designed to help students understand the real world applications of the material taught.

Both types of assignments concentrate on narrowly defined topics and are similar to traditional homework assignments used in non-web based courses. As the course progresses, the assignments are announced, handouts are distributed, and the assignments are previewed in class. At this time, the assignment is also made available on the course web site. The online assignment can be completed up to the due date. After the due date, the score is discounted by twenty five percent per day. Assignments can be revisited without penalty throughout the remainder of the semester. Each assignment typically consists of around ten problems requiring numerical answers and menu selections.

The online homework assignments are instantly scored when a “Check Answers” button is clicked. This “Check Answers” procedure is designed to let students’ monitor their progress. Each student has the same skeleton problem for each assignment but variations in assumptions give each student a unique set of correct answers.

Student interactions with the homework assignments are recorded on the web server. Each page load or reload by each user is recorded in an access log. Data recorded in the access log includes the user’s ID, name, current date and time (timestamp), and page loaded, and result (e.g. login, log-out, access denied, or assignment loaded / reloaded). This file contains a time-stamped record for each “Check Answer” click for every student and for every homework assignment. These data are filtered by comparison to the course calendar; site hit types, and student identity.

As students work, they can save their work by clicking a “Save / Submit” button. This action records the timestamp, user identity, homework identifier, response to each question, and overall score in the “all Submits” data table. All Submits thus contains a record for each click of the “Save / Submit” button.

The “Last Submit” table contains only one record per assignment and per user. This table records students’ final assignment scores and responses as reflected in the course gradebook. The “Last Submit” table also contains raw scores and discounted scores. These two scores differ only for homework submitted after the due date.

We mainly use the final submit scores from each assignment. We not only look at the scores that contribute to students' learning outcomes for specific topics, but also try to explore certain activities that relate to preferred learning styles. These data will be examined to determine their relationship with individual test item responses. We also study the distribution of submit timestamps purely on the time-axis.

Class attendance measures of participation in the traditional aspects of this course. Pop quizzes are used in this course and are called "Bonus points" because they are added on to students' class scores after grade boundaries have been determined. Bonus points indicate attendance because non-zero (though potentially very small) scores are awarded for any bonus points submitted. Approximately fifteen pop quizzes are given over the course of a semester. We use the number of non-zero bonus points earned by a student as an estimator of attendance.

The course gradebook contains the students' scores as well as other information such as major and academic classification and scores on homeworks, exams and bonus points.

We use examination scores as measurements of learning outcomes. This course utilizes two midterms and a final exam. The exam format is multiple-choice with each question focused on a narrow topic. The midterms each contain 25 questions and final contains 30 questions. Each exam is administered in two forms with differing choice orderings. Student seating is assigned at exam time and each student identifies the form of the exam completed. Exams difficulties are consistent across years.

Students' responses to each question are recorded in the grade book, allowing analysis of responses to each question.

Each exam question derives from a narrowly defined topic taught in the course. Exam item topics are frequently tied to homeworks or bonus points so we can relate homework practice to related examination item proficiency. However finding quantitative evidence that in class presentation influences learning will be difficult because a single lecture can cover many narrow topics. In this sense, it is easier to relate exam question topics to assignment topics rather to presentation topics. Frequently homework questions can be found in exams with slight changes in assumptions. This strong relationship between exam questions and assignment questions permits linking exam questions to corresponding homeworks. This pairing allows us to consider the linkage between practice and learning in the course. The detail of matching homework items to exam items will be discussed in the following sections.

The course also includes an online course evaluation to capture students' feedback at the completion of the course. Although these data are primarily collected to improve the course design, instruction, and management, information about the student is also collected. This evaluation is given out online at the end of the semester. Students can receive one point towards the final grade (105 points possible) for completing the questionnaire. The responses are collected online.

While the evaluation questionnaire is composed of six sections, our study mainly uses two sets of information – the learners' characteristics and Learning Style Inventory (LSI) scores. The data collected from the course evaluation are self-reported. Although self-reporting creates possibilities for errors the data provide valuable details about students' preferred learning styles and characteristics.

The Learning Style Inventory is an instrument that measures students' preferences for specific learning styles. A learning style is a student's consistent way of reacting to provided learning materials. Keefe (1979) defines learning styles as the "composite of characteristic cognitive, affective, and physiological factors that serve as relatively stable indicators of how a learner perceives, interacts with, and responds to the learning environment."

The LSI, asks students to rank nine sets of four descriptive terms arranged in a table. Each of the four columns corresponds to a learning style category. Those categories are Concrete Experience, Reflective Observation, Abstract Conceptualization and Active Experimentation. The Concrete Experience (CE) column represents students' preferences for learning by doing, the Reflective Observation (RO) column represents preferences for learning by watching, the Abstract Conceptualization (AC) column represents preferences for learning by abstract thinking, and the Active Experimentation column represents learning by experimentation.

3.2.3 Databases and their contents

Table 2 summarizes the datasets and variables used in this study.

Table 2. Description of Dataset Variables

Dataset	Variable	Description
Grade Book		Database with one record per student where scores for all graded course components are stored. Also contains academic related demographic data such as class, major, and college.
	Student ID	Student identifier, i
	$ScoreEx_{i,x}$	Student i's score on exam x(x=1,2,3; Midterm1, Midterm2, Final Exam)
	$ScoreItem_{i,x,y}$	Student i's response to item y on exam x. By comparing to answer key, can be converted to score (1 if correct answer selected, 0 otherwise)
	$ScoreTotal_{i,k}$	Student i's final score on homework k
	$BP_{i,k}$	Student i's score on popquiz k, as the bonus point. If student was present then a score greater than 0 was recorded, regardless of the quality of the response.
AccessLog		Text file that records all pages loaded and reloaded (triggered by check answers click) requested from the course web site.
	Timestamp	Date and time of page load
	Student ID	Student identifier, i
	Homework ID	Homework that the loaded webpage is represented, k
	Page ID	Identifier indicating webpage loaded
	Access Type	Type of access resulted from a click
Course Evaluation		Database of student responses to an online course evaluation survey. One record per student respondent.
	Student ID	Student identifier, i
	$CGPA_i$	Student i's reported cumulative grade point average
	$HrsWorked_i$	Student i's reported working hours per week
	$CreditsperSem_i$	Student i's registered course credits per semester
	$Gender_i$	Student i's reported gender, 1 for male
	$USWhite_i$	Student i's reported demographic group, 1 for US-white
	$USMinority_i$	Student i's reported demographic group, 1 for US-minority
	$Sophm_i$	Student i's reported demographic group, 1 for sophomore
	$Junior_i$	Student i's reported demographic group, 1 for junior
	$Senior_i$	Student i's reported demographic group, 1 for senior
	$LSI_{i,l}$	Student i's reported learning style ranking for column l
HWSubmits		Database of homework submittal scores
	Homework ID	Homework identifier, k
	Student ID	Student identifier, i
	Timestamp	Date and time of homework page load, as in AccessLog
	$ScoreTotal_{i,k}$	Student i's final score on homework k
	$HitsOnline_{i,k}$	Student i's number of hit sites contributed to homework k

3.2.4 Data manipulation

1. Definition of active learning sessions.

Online behavior is recorded as students' logons, log outs and clicking activity involving the "Save / Submit" and "Check Answers" buttons. The student's learning effort is quantified by the sum of the time segments recorded in the applicable assignment period. This is a coarse measurement of learning time in that it does not necessary reflect "active" learning because a student can log-on and submit a partial answer before s/he walks away to have a cup of coffee or goes to bed before returning to the assignment. Based solely on time stamps, coffee breaks and sleep time would be counted as learning effort. We observe some extremely long sessions lasting for days or months.

To remedy this, we define "sessions" as active learning time that includes only frequent site hits, and turn off the clock after an inactivity time threshold has been reached. In other words, site hits are only considered as part of active learning "sessions" if the elapsed time since the previous hit is less than the threshold.

The empirical issue is what should be the time threshold that delineates an active learning session. The time threshold for an active session is established by examining the data. This threshold should be neither too short as that would result in the loss of valuable information nor too long as that would cause sessions to include non-learning time. The criterion used is that the threshold is considered representative if the portion of time intervals longer that the session threshold is a negligible portion of all intervals considered. Figure 1 shows the distribution of all time intervals between site hits on all course materials by all students in a year.

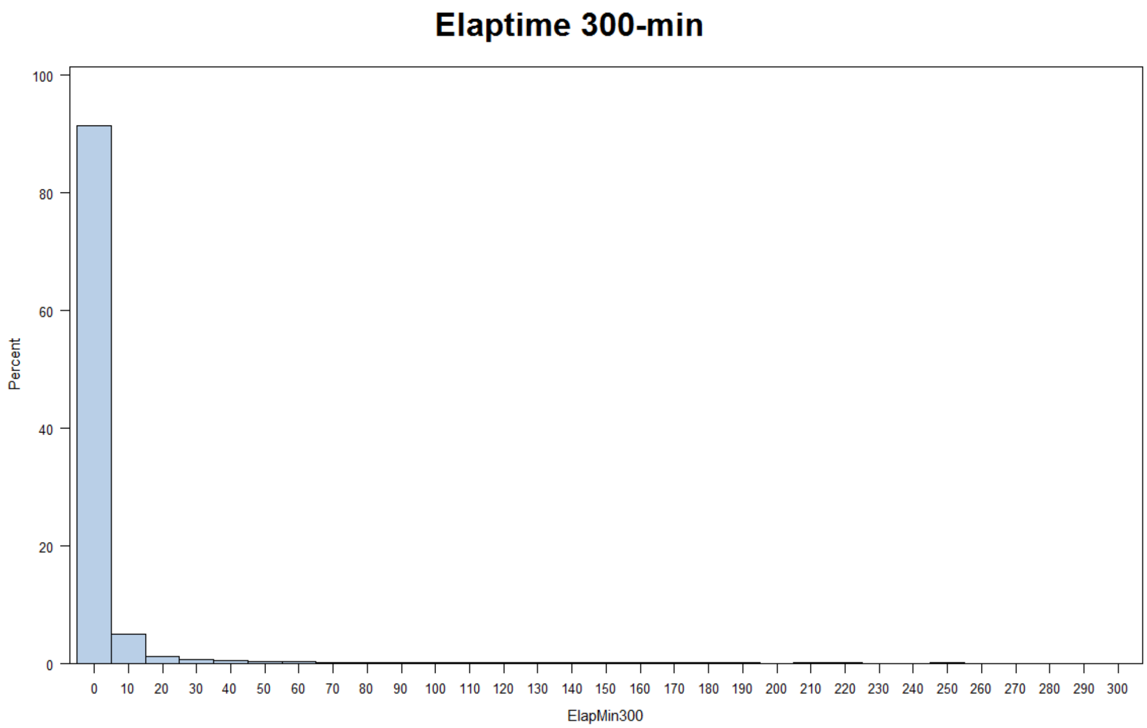

Figure 1. Distribution of Time Intervals between Site Hits

This figure indicates that more than 95% of the site hit time intervals are less than 30 minutes. So we regard site hit intervals of more than 30 minutes to contain substantial non-academic activities. Those sessions between 0 and 30 minutes, are considered to represent learning efforts, and are summed and represented by the *OnlineTime* variable in our regressions.

2. Matching examination items to homework items by topic.

We “pair” exam items with online homework items when the examination topic matches an online homework topic. Twenty–five such pairs are identified in table 3.

Table 3. Pairings of Homework and Examination Items

	Homework Items	Exam Items
1	HW2, all parts	Q7@Midterm I
2	HW2, all parts	Q8@Midterm I
3	HW3	Q15@Midterm I
4	HW3, part2 and part5	Q17@Midterm I
5	HW4, all parts	Q23@Midterm I
6	HW5, given P find disc yld	Q11@Midterm II
7	HW5, 2 yr T bond, find price	Q12@Midterm II
8	HW6, parts 2, 3, 4	Q16@Midterm II
9	HW6, part 6	Q20@Midterm II
10	HW7, part 1	Q23@Midterm II
11	HW7, part 5. FTE5	Q24@Midterm II
12	FTE4	Q21@Midterm II
13	FT5	Q24@Midterm II
14	HW3, parts 1, 4	Q6@Final
15	HW4	Q11@Final
16	HW5	Q15@Final
17	HW6, part 1	Q16@Final
18	HW6, part 7.	Q19@Final
19	HW7, part 5.	Q21@Final
20	HW8	Q24@Final
21	FTE6	Q11@Final
22	FTE6	Q19@Final
23	FTE5	Q21@Final
24	FTE6	Q26@Final
25	FTE6	Q27@Final

As can be seen from this table, an exam question might relate to more than one homework question but examination questions never relate to more than one homework.

3. Learning Style Inventory scoring

As described above, the four columns of the learning style inventory table correspond to each of the preferred learning styles: CE (concrete experience), RO (reflective observation), AC (abstract conceptualization), and AE (active experimentation). Because of the required rankings (one through four) in each row the sums for each of the four columns are not independent. Also, RO and AE, and CE and AC are pairs of opposite preferences, so the four correlated variables can be reduced to two. We compute the two composite scores, (AC-CE) and (AE-RO). These two variables indicate the preference for a learning style vis-à-vis it polar opposite. For example, a large negative (AC-CE) indicates that concrete experiences is more strongly preferred over abstract conceptualizations and a strongly negative (AE-RO) indicates that reflective observation is more strongly preferred over active experimentation.

3.2.5 Variable characteristics

Table 4 presents the summary statistics for the variables used in our analysis. The variables include students' own reported information, their homework and test data, and the learning style data.

Table 4. Descriptive Statistics for Data Used

	Summary of Responses			
Variable	Variable Definition	N	Avg	Std.Dev
$CGPA_i$	Cumulative grade point average	84	2.8676	0.5134
$CreditsperSem_i$	Credits completed during semester	84	13.5595	2.5429
$HrsWorked_i$	Hours per week, on average, employed during semester	84	14.3690	12.4579
$ClsParticip_i$	Total Bonus Gains = $\Sigma BP_{i,k}$	84	11.1905	4.0679
$Gender_i$	Dummy variable: 1 if male student, 0 otherwise	84	0.8333	0.3749
$USWhite_i$	Dummy variable: 1 if U.S. white student, 0 otherwise	84	0.6786	0.4698
$USMinority_i$	Dummy variable: 1 if U.S. minority student, 0 otherwise	84	0.2381	0.4285
$Sophm_i$	Dummy variable: 1 if sophomore student, 0 otherwise	84	0.0833	0.2780
$Junior_i$	Dummy variable: 1 if junior student, 0 otherwise	84	0.3690	0.4854
$Senior_i$	Dummy variable: 1 if senior student, 0 otherwise	84	0.5357	0.5017
$ScoreTotal_{i,k}$	Assignment score in the "last submit"	1320	21.4800	24.3100
$ScoreOnline_{i,k}$	Standardized Assignment score in the "last submit"	1320	0.7716	0.3663
$PointsTotal_i$	Total score for all assignments = $\Sigma ScoreTotal_{i,k}$	84	325.38	66.9094
$TimeOnline_{i,k}$	Number of hours worked on each assignment=$\Sigma(\Delta Timestamp)$ for all sessions	1320	0.8236	0.9366
$HitsOnline_{i,k}$	Number of total site hits each assignment	1320	34.5296	38.0171
$HitsTotal_i$	Number of total site hits for all assignment= $\Sigma HitsOnline_{i,k}$	84	768.0238	411.5462
$Active_i$	Dummy variable of course participation activity; 1 if completes all online assignments, 0 otherwise.	84	0.4286	0.4978

Table 4. Descriptive Statistics for Variables Used (Cont'd)

	Summary of Responses			
Variable	Variable Definition	N	Avg	Std.Dev
Course Examination Data				
$ScoreEx_{i,1}$	Number of correctly answered midterm1 questions	84	17.1071	3.9240
$ScoreEx_{i,2}$	Number of correctly answered midterm2 questions	84	16.2458	4.6270
$ScoreEx_{i,3}$	Number of correctly answered final exam questions	84	16.7619	4.2273
$ScoreItem_{i,x,y}$	Correctness of the answer to an exam question	2100	0.4976	0.3726
Composite Learning Style Inventory Scores				
CE_i	Concrete experience (preference for learning by doing)=$LSI_{i,0}$	84	16.1667	3.5562
RO_i	Reflective observation (preference for learning by watching=$LSI_{i,1}$	84	17.4881	3.0080
AC_i	Abstract conceptualization (preference for learning by abstract thinking=$LSI_{i,2}$	84	18.9762	3.1354
AE_i	Active experimentation (preference for learning by experimentation=$LSI_{i,3}$	84	17.4762	3.4969
$LSI20_i$	Independent Learning Style Indicator 1 = $LSI_{i,2}$ - $LSI_{i,0}$	84	2.8095	5.8340
$LSI31_i$	Independent Learning Style Indicator 1 = $LSI_{i,3}$ - $LSI_{i,1}$	84	-0.0119	5.6323

CHAPTER 4 EMPIRICAL MODELS AND EXPECTED RESULTS

4.1 Online Learning Effort and Outcomes

We will use the data described in the previous chapter to address questions of importance in delivering web-based courses. These questions include: is our online approach effectively delivering and evaluating course concepts? Is students' mastery of knowledge increased by using online delivery methods? Does testing reinforce learning?

The answers to these questions require tying student's efforts (input) to learning outcomes (output). If the effort-outcome relationship for online learning is positive and significant, then online teaching tools are effective. If not, then other justifications such as efficiency in allocation of teaching resources merit further investigation.

Homework scores and exam scores both serve as output measures for online learning outcomes. We will analyze both types of data. Online homework scores will be viewed as an intermediate output in the production of test scores. We first analyze how online homework scores are produced. Next we will match these data to test score data to determine whether online homework contributes to learning as measured by test scores. This analysis will require topic-wise pairing of test items with online homework assignments. Pairing online learning efforts with corresponding test item scores depicts the relationship between online efforts and learning more accurately than simply observing aggregate exam scores.

4.2 Online Homework as an Intermediate Input to Learning

Online homework is an intermediate input in the production of learning. A model of how this intermediate form of learning contributes to overall learning has been developed.

The model uses recorded online behavior of each learner on each homework assignment as the observational unit. It is represented as:

$$\begin{aligned}
ScoreOnline_{i,k} = {} & \alpha + \beta_0 HrsWorked_i + \beta_1 CreditsperSem_i + \beta_2 ClsParticip_i \\
& + \beta_3 TimeOnline_{i,k} + \beta_4 HitsOnline_i + \beta_5 CGPA_i + \beta_6 Gender_i \\
& + \beta_7 USWhite_i + \beta_8 USMinority_i + \beta_9 Sophm_i + \beta_{10} Junior_i \\
& + \beta_{11} Senior_i + \sum_{k=1}^{14} \gamma_k AssignmentDummy_{i,k} \\
& + \sum_{k=1}^{14} \delta_k TimeOnline_{i,k} \times AssignmentDummy_{i,k} \\
& + \lambda_1 ScoreOnline_{i,k-1} \\
& + \varepsilon_{i,k}
\end{aligned} \tag{4.1}$$

Where $ScoreOnline_{i,k}$ is student *i*'s score on homework *k* as recorded in the course grade book at the end of the semester. These scores are unitized as the proportion of possible (between 0 and 1) in order to make assignments comparable. Despite this unitization, we still need to account for distribution differences among assignments, due to varying difficulty levels within and across the assignments. To account for these differences, we standardize the unitized homework scores using each assignment's mean and standard deviation.

$HrsWorked_i$ and $CreditsperSem_i$ are self-reported variables which represent the number of hours that student *i* worked per week, and the number of credits attempted by student i in the semester. $ClsParticip_i$ represents class attendance measured as student i's presence when pop-quizzes were administered.

$TimeOnline_{i,k}$ represents the time(in hours) student *i* spent on assignment *k.* In constructing this measure we excluded the observations that reflect unrealistically lengthy access times under the assumption that the student could load a page then turn attention elsewhere. $HitsOnline_{i,k}$ represents the number of webpage loads/reloads of assignment *k*.

Page reloads reflect confirmation-seeking obtained by clicking the “Check Answers” button on the homework page. This variable reflects the effort expended by the student in working with the online materials. It also represents learning styles as we would expect that those that prefer to learn from active experimentation would hit the site through the “Check Answers” click more frequently than students with other learning style preferences.

$CGPA_i$ represents student *i'*s cumulative grade point average. Dummy variables include $Gender_i$ (Male=1, Female=0), $USWhite_i$, $USMinority_i$ (Non-US excluded to avoid linear dependency), and $Sophm_i$, $Junior_i$ and $Senior_i$ to represent student *i*’s academic classification (freshmen excluded to avoid linear dependency).

Assignments are not equal in length or difficulty and on average require differing time allocations to achieve an average score. To distinguish the effect of efforts spent on different homework assignments, we add dummy variables $AssignmentDummy_k$ for assignment k=1 to 14. This will account for differences among the assignments. The interaction of assignment dummies and $TimeOnline_{i,k}$ measures the marginal productivity of online time in working assignment k.

We also included the previous standardized proportional homework score $ScoreOnline_{i,k-1}$ in each regression, in order to account for time allocation strategies exercised dynamically by student i. OLS estimates of the parameters in equation (4.1) are presented in Table 5.

Table 5. Regression Estimates for Standardized Proportional Online Homework Score[1]

Model	Baseline		Baseline + HW.Dummies		Baseline + HW.Dummies+ HW.Slopes		Baseline + HW.Dummies+ HW.Slopes+ Previous HWs	
Variables	Est.Coff	Pr>|t|	Est.Coff	Pr>|t|	Est.Coff	Pr>|t|	Est.Coff	Pr>|t|
Intercept	-0.1647	0.4283	-0.2747	0.1803	-0.2794	0.1713	-0.1639	0.4034
TimeOnline	0.1402	0.0002	0.3241	<.0001	0.3430	0.0021	0.2496	0.0194
HitsOnline	0.2545	<.0001	0.2480	<.0001	0.2466	<.0001	0.2058	<.0001
CGPA	0.1428	<.0001	0.1222	<.0001	0.1218	<.0001	0.0745	0.0022
HrsWorked	0.0612	0.0308	0.0695	0.0066	0.0720	0.0046	0.0523	0.0320
ClsParticip	0.2614	<.0001	0.2502	<.0001	0.2486	<.0001	0.1648	<.0001
CreditsperSem	0.0171	0.5412	0.021	0.4070	0.0263	0.2955	0.0208	0.3874
Gender/Male	-0.1364	0.0470	-0.1378	0.0262	-0.1389	0.0239	-0.1017	0.0843
USWhite	0.0103	0.9171	0.0738	0.4076	0.0869	0.3289	0.1000	0.2401
USMinority	0.1291	0.2395	0.1813	0.0675	0.1898	0.0544	0.1576	0.0954
Sophm	-0.1054	0.6001	-0.1938	0.2858	-0.2142	0.2341	-0.2018	0.2416
Junior	0.1713	0.3461	0.1128	0.4921	0.1108	0.4962	0.0346	0.8248
Senior	0.2447	0.1702	0.1897	0.2390	0.1873	0.2414	0.0967	0.5282

1. N=1320 for first 3 models, and N=1176 for the 4th model)

Table 5. Regression Estimates for Standardized Proportional Online Homework Score (Cont'd)

Model	Baseline		Baseline + HW.Dummies		Baseline + HW.Dummies+ HW.Slopes		Baseline + HW.Dummies+ HW.Slopes+ Previous HWs	
Variables	Est.Coff	Pr>|t|	Est.Coff	Pr>|t|	Est.Coff	Pr>|t|	Est.Coff	Pr>|t|
FTE0.Dummy			0.5218	<.0001	0.6181	<.0001	1.1575	<.0001
FTE1.Dummy			0.4571	0.0002	0.5856	<.0001	0.6912	<.0001
FTE2.Dummy			0.4663	0.0001	0.5584	<.0001	0.4121	0.0041
FTE3.Dummy			0.2757	0.0226	0.2772	0.0210	0.0859	0.4617
FTE4.Dummy			-0.1389	0.2533	-0.1808	0.1372	-0.3769	0.0015
FTE5.Dummy			0.0703	0.5595	0.0681	0.5690	-0.0431	0.7095
FTE6.Dummy			-0.0406	0.7364	-0.0427	0.7209	-0.1329	0.2471
HW2.Dummy			0.5388	<.0001	0.5419	<.0001	0.4125	0.0006
HW3.Dummy			0.1813	0.1323	0.1807	0.1306	-0.0009	0.9941
HW4.Dummy			0.5870	<.0001	0.6606	<.0001	0.4009	0.0018
HW5.Dummy			-0.0692	0.5742	0.0337	0.8027	-0.1397	0.2837
HW6.Dummy			-1.0012	<.0001	-0.8064	<.0001	-0.9726	<.0001
HW7.Dummy			-0.4132	0.0007	-0.4544	0.0002	-0.4954	<.0001

Table 5. Regression Estimates for Standardized Proportional Online Homework Score (Cont'd)

Model	**Baseline**		**Baseline + HW.Dummies**		**Baseline + HW.Dummies+ HW.Slopes**		**Baseline + HW.Dummies+ HW.Slopes+ Previous HWs**	
Variables	Est.Coff	Pr>|t|	Est.Coff	Pr>|t|	Est.Coff	Pr>|t|	Est.Coff	Pr>|t|
FTE2.Dummy *TimeOnline					0.0263	0.1276	0.3612	0.0536
FTE3.Dummy *TimeOnline					0.0720	0.5290	-0.0162	0.9045
FTE4.Dummy *TimeOnline					0.2486	0.1325	0.2619	0.0497
FTE5.Dummy *TimeOnline					0.3430	0.9188	0.0347	0.8049
FTE6.Dummy *TimeOnline					0.2466	0.9810	0.0754	0.5620
HW2.Dummy *TimeOnline					-0.1389	0.9864	0.0823	0.6307
HW3.Dummy *TimeOnline					0.1898	0.9589	0.0754	0.6327
HW4.Dummy *TimeOnline					0.0869	0.2534	0.1768	0.3213
HW5.Dummy *TimeOnline					-0.2142	0.2179	-0.1176	0.3657
HW6.Dummy *TimeOnline					0.1108	0.1374	-0.0652	0.5523
HW7.Dummy *TimeOnline					0.1873	0.3162	0.1722	0.1713
1st Previous Homework Score							0.2277	<.0001
R-Square[1]	0.2725	<.0001	0.4135	<.0001	0.4317	<.0001	0.4808	<.0001

1. The probability is the Pr > F

This table shows that the variable $TimeOnline_{i,k}$ is significant ($p<0.05$) across all models, meaning that student time spent on web based homework assignments is positively and significantly related to the assignment score. These results indicate that for each hour a student spends on online assignments, s/he gains 3% of the standardized possible homework score $d\,(HW_{ik}/\,se_k)\,/\,dTimeOnline_{ik} = 0.03$ so $dHW_{ik}/dTimeOnline_{ik} = 0.03\ se_k$. The approach represented by the "Check Answers" button is a useful learning tool as these clicks, measured by $HitsOnline_{i,k}$, are positively and significantly associated with homework scores.

The cumulative GPA measures students' academic proficiency and is positively correlated with the scores ($p<0.005$). It is more significant than the online time variable.

$HrsWorked_i$ positively and significantly influences homework scores ($p<0.01$). Our initial expectation was that this relationship would be negative. This discrepancy might be due to attributes of online courses that facilitate learning for time and place challenged students. As opposed to traditional methods, online methods provide more flexibility and let working students study at their convenience. This allows better time management, which works to the comparative advantage of working students. Although statistically significant, the significance of the *HrsWorked* variable is not as strong as the other factors, it does reveal some advantages of online courses.

$ClsParticip_i$ also contributes significantly to the homework score outcomes indicating that the time that students spent in traditional classroom learning is important in generating good learning outcomes. In comparing the magnitude of the coefficients, note that each class participation unit is 50 minutes of class time while the time spent on online assignments is measured in hours. Learners' "attitudes" may also influence this relationship as generally harder-working learners attend classes more frequently and gain better scores on homeworks and tests. Furthermore, it reveals the fact that traditional teaching and learning approaches complement web-facilitated teaching and learning.

The *Gender* variable indicates that males usually do worse than females for a given amount of time spent on homework, and US minority groups usually do better with online course delivery, again assuming equal time investments. We don't observe significant effects for different student classes, nor does the number of credits attempted seem to influence online homework scores.

The homework dummies reflect the variability in the complexity of the problem sets, e.g. FTE0, FTE1, FTE2, FTE3, FTE4, HW2, HW3, and HW4 have higher average scores than HW9, and HW6 and HW7. The interaction term coefficients estimate the marginal benefit of time spent on individual assignments. That is to say, efforts spent on FTE0, FTE1, FTE3, HW5 and HW8 have significant and positive marginal productivity.

The influences of previous homework scores are also positive and significant.

4.3 Online Homework as a Direct Contributor to Learning

Traditional education methods have some advantages over computer assisted technologies. These advantages include instant interaction, class monitoring, and real-time feedback. Hybrid courses exploit the relative advantages of both methods. Given the hybrid nature of this course, looking at the effects of the online components without considering the total product gives an incomplete analysis of web-facilitated education.

In the previous section, we found a significant relationship between online homework effort and homework scores. We now investigate the connection between those online learning materials and learning. Here we map homework outcomes to a larger outcome which is the students' understanding of the course.

Hence, the issue is how to evaluate the contribution of the online component relative to overall learning? Further, since learning is a long-term phenomenon, estimating its short run magnitude will inevitably be inaccurate. Nonetheless, we measure each individual's learning outcomes by test scores. We evaluate the contribution of online homework to learning by regressing homework scores against examination items on the corresponding topic. This is accomplished by matching examination item topics to the online assignments.

Online assignments are designed to provide learners with practice problems. A variety of questions are offered to consolidate students' understanding of course materials. Thought is elicited on homework assignments with the use of multiple choice, fill in the blank, and check applicable completion statements while examinations use the multiple choice format exclusively. Although the assignments and tests cover the same topics, the response formats differ, adding to the difficulty of comparing responses. The exam-question / homework assignment matching process was discussed in chapter 3 where we identified 25 topic-linked pairs.

We depict the impact of online homework on overall learning with the probit model:

$$\Pr(ScoreItem_{i,j} = 1 \mid ScoreOnline_{i,j}, CGPA_i) = \Phi[\alpha ScoreOnline_{i,j} + \beta CGPA_i] \tag{4.2}$$

where $ScoreItem_{i,j}$ equals 1 for a correct response by student i on exam question-homework assignment pair j *(0 otherwise)*; $ScoreOnline_{i,j}$ is student i's score on the homework assignment that is paired (j) with the exam question, and $CGPA_i$ is student i's cumulative grade point average.

Exam response correctness takes a value of either 0 or 1. The expected probability of a correct answer is represented with a probit model with the variable representing the learning attributable to online homework measured by the homework score on the corresponding assignment. The student's cumulative grade point average is included to represent the student's learning proficiency.

We expect to observe that engaging in homework will have a positive effect on examination performance on topics covered by the homework. If so, then we have evidence that online practice positively affects the measured learning outcome.

For this analysis we identified 25 examination items covering concepts that are also covered by the online homework assignments. Each student faced these pairings. Complete data were available for 84 students resulting in a total of 2100 observations.

We standardized each homework assignment score to account for the differing degrees of complexity within the assignments and utilized dummy variables to account for differences in difficulty among the examination items. The results of the analysis are shown in Table 6.

In table 6 shows that homework scores and CGPA significantly and positively influence the probability of a correct response on the related examination item. The cumulative grade point average displays the expected positive relationship with exam outcomes. We note that homework scores have an even more significant effect on examination performance than CGPAs. From these results we can infer that doing well on the online homework is at least as effective as high CGPAs in generating correctly-

Table 6. Estimated Regression Coefficients for Homeworks' Contribution to Final Learning [1]

Variable	Est.Coff	Std.Err	t-value	Pr>|t|
Intercept	0.5390	0.2635	3.86	0.0002
CGPA	0.0968***	0.0433	3.39	0.0012
ScoreOnline	0.1002***	0.0135	3.56	0.0008
PAIR1	-0.5127**	0.1497	-2.63	0.0102
PAIR2	-1.1145***	0.2030	-13.97	<.0001
PAIR3	-0.8514***	0.1978	-9.45	<.0001
PAIR4	-0.7869***	0.1042	-8.07	<.0001
PAIR5	-0.7555***	0.1545	-3.83	0.0002
PAIR6	-0.4821*	0.1356	-2.49	0.0157
PAIR7	-0.9775***	0.0112	-10.13	<.0001
PAIR8	-0.8147***	0.2084	-8.85	<.0001
PAIR9	-0.7250***	0.1674	-3.37	0.0003
PAIR10	0.2975	0.1903	1.41	0.1625
PAIR11	0.0308	0.1353	0.22	0.8809
PAIR12	-1.7855***	0.1488	-15.63	<.0001
PAIR13	0.1477	0.1387	2.03	0.4782
PAIR14	-1.7201***	0.1849	-15.10	<.0001
PAIR15	0.7193**	0.1630	3.17	0.0022
PAIR16	-0.5126*	0.2029	-2.63	0.0102
PAIR17	-0.7860***	0.1685	-8.08	<.0001
PAIR18	-0.2967††	0.0936	1.49	0.1393
PAIR19	-1.3011***	0.1774	-14.33	<.0001
PAIR20	-1.5559***	0.1640	-14.61	<.0001
PAIR21	-0.8496***	0.1400	-9.05	<.0001
PAIR22	0.0322	0.2920	0.13	0.8755
PAIR23	-0.7874***	0.1336	-8.22	<.0001
PAIR24	-0.7555***	0.1287	-3.85	0.0002

1. N=2100, (Pr>F) <0.0001

answered exam questions. If correctly responding to examination items represents learning, then these results match our initial expectations.

We have observed that online problem sets generally translate to higher test scores. The online homework methodology helps students grasp important and difficult points in the course. Hence, student efforts on online homework lead to increased learning outcomes.

4.4 Learning Styles and Online Homework Participation

Thus far, we have considered online homework's contribution to learning. We now turn our attention to how educational technology might discourage learning. Compared to traditional education, online education suffers from higher student dropout usually preceded by nonparticipation. This phenomenon is a concern to educators, and is due to the flexible nature of the online format. What causes learners' non-participation in courses and frequently leads to a drop-out decision? Is it related to students' preferred learning styles? What learning styles contribute to non-participation? Insights into these complex and hidden aspects of online learning are hampered by a lack of data. Although some explicit student characteristics such as gender, ethnicity, and individuals' time allocation patterns have been identified, factors such as learners' personality, economic status, and major life events might also be involved. While the root cause might be multi-facetted, our study focuses on students' learning style preferences.

The phenomenon that catches our interest in this hybrid web-learning course is that many students leave homework unfinished. This phenomenon includes submittal of incomplete assignments, discontinuation of effort as the course progresses, and failure to meet submission deadlines.

4.4.1 Determinants of Active Participation

Our analysis of this issue will proceed as follows. First, we assign our observations to two groups – designated as active participants and passive participants – based on their completion of online assignments. Students who complete all assignments and course evaluations are assigned to the former group. The latter group includes students who have only partially completed the required online assignments and the final evaluation form.

We define $Active_i$ as a participation variable that distinguishes between students who complete all homework assignments and those who don't. $Active_i = 1$ if student i completes all homework, while $Active_i = 0$ otherwise. We use $Active_i$ as a dependent variable and investigate whether it is influenced by individual characteristics or learning styles. Explanatory variables are group dummies such as gender, academic class level, demographic characteristics, and measured learning style preferences. We use the probit model to estimate the relationship:

$$\begin{aligned}\Pr(Active_i = 1 \mid X_i) = \Phi[&\alpha + \beta_0 HrsWorked_i + \beta_1 CreditsperSem_i + \beta_2 ClsParticip_i \\ &+ \beta_4 CGPA_i + \beta_5 Gender_i + \beta_6 USWhite_i + \beta_7 USMinority_i \\ &+ \beta_8 Sophm_i + \beta_9 Junior_i + \beta_{10} Senior_i + \gamma_1 LSI20_i + \gamma_2 LSI31_i] \end{aligned} \tag{4.3}$$

where $HrsWorked_i$, $CreditsperSem_i$, $ClsParticip_i$, $CGPA_i$, $Gender_i$, $USWhite_i$, $USMinority_i$, $Soph_i$, $Junior_i$, $Senior_i$ are as defined above and $LSI20_i$ and $LSI31_i$ are defined in Table 4 of Chapter 3.

We then test $H_0 : [\beta_0 ... \beta_{10} \gamma_1 \gamma_2] = \underline{0}$ If the null hypothesis is rejected, then it means that participation is influenced by students' characteristics and learning styles.

From table 7, we see that online participation behavior is significantly and positively related to traditional face-to-face class participation (p=0.0002). Females and non-US students tend to participate more than their peers by completing homeworks. We also observe that the learning style indicator LSI31, which represents the difference between active experimentation and reflective observation, contributes to being an active

Table 7. Probit Regression Estimates for Determinants of Participation[1]

Variable	Est. Coff	Std. Err	χ^2-value	Pr>χ^2
Intercept	1.1900	0.2046	0	0.9954
CGPA	0.3507	0.3531	0.9863	0.3206
Credits	0.0378	0.0769	0.2416	0.6231
HrsWorked	0.0151	0.0147	1.0508	0.3053
ClsParticip	0.2271***	0.0602	14.2175	0.0002
Gender/Male	-0.7899†	0.4661	2.8724	0.0901
USMinority	-1.2174††	0.8049	2.2875	0.1304
USWhite	-1.3757†	0.7475	3.3872	0.0657
Sophm	-4.3719	0.2046	0.0005	0.9829
Junior	-3.9277	0.2046	0.0004	0.9847
Senior	-3.7724	0.2046	0.0003	0.9853
LSI20	0.0029	0.0327	0.0079	0.9292
LSI31	0.0756*	0.0357	4.4928	0.0340

1. N=84, Pr > χ^2 = 0.0006.

participant. Other group characteristics can be summarized as: US minority groups and US white are more passive than international students.

4.4.2 Determinants of Site-hitting Behavior

An alternative measure of participation is activity frequency which may be influenced by learning style preferences. $HitsOnline_{i,k}$, the number of hits recorded by student i in doing online assignment k, measures the student's level of active involvement in reaching a target score. $HitsOnline_{i,k}$ has more granularity than the $Active_i$ variable.

We conduct a regression to determine whether $HitsOnline_{i,k}$ is influenced by demographic group or learning styles. We regress $HitsOnline_{i,k}$ on the group identification variables, the LSI scores of individual learners and the site-hitting behavior on the previous assignment. Lagged hits are used to capture discontinuation of effort effects.

We develop the following model:

$$\begin{aligned} HitsOnline_{i,k} = {} & \alpha + \beta_0 HrsWorked_i + \beta_1 CreditsperSem_i + \beta_2 ClsParticip_i + \beta_5 CGPA_i \\ & + \beta_6 Gender_i + \beta_7 USWhite_i + \beta_8 USMinority_i + \beta_9 Sophm_i + \beta_{10} Junior_i + \beta_{11} Senior_i \\ & + \beta_{12} LSI20_i + \beta_{13} LSI31_i + \beta_{14} Active_i \\ & + [\sum_{k=1}^{14} \gamma_k AssignmentDummy_{i,k}] + \sum_{j=1}^{11} \lambda_j HitsOnline_{i,k-j} + \varepsilon_{i,k} \end{aligned} \tag{4.4}$$

and test $H_0 : [\beta_0 ... \beta_{14} \gamma_1 ... \gamma_{14} \lambda_1 ... \lambda_3] = \underline{0}$, where *HrsWorked*$_i$, *CreditsperSem*$_i$, *ClsParticip*$_i$, *CGPA*$_i$, *Gender*$_i$, *USWhite*$_i$, *USMinority*$_i$, *Soph*$_i$, *Junior*$_i$, *Senior*$_i$ *LSI20*$_{i,}$ *LSI31*$_i$, *Active*$_i$, *AssignmentDummy*$_i$ are as defined above, and *HitsOnline*$_{i,k-j}$ is the number of hits on jth assignment prior to assignment k.

Site hits varies by assignment due to differences in the length and complexity of the assignment. It is standardized for each assignment. After doing this, each assignment will have the same mean (0) and variance (1). This standardization eliminates the need for the dummy variables above and creates a homoscedastic dependent variable.

HitsOnline$_{i,k-j}$ is also standardized and is the site hits for student *i* on the previous assignment. Note HW and FTES are intermingled and we sort the two types of assignments according to due dates. These lagged effects account for ‘quitting’ or less engaged site-clicking behavior for online assignments. From table 8, we see that the site hits is largely determined by the student’s group characteristics and previous site hits. A strong temporal effect is observed as participation in the most recent assignment strongly and significantly influences participation in the current assignment. This reflects the discontinuation phenomenon in which some students reduce the intensity of their online assignment activities until they cease to show progress. For the group-wise explanatory variables, males tend to click “Check Answers” more than females, and non-native students click “Check Answers” more than US natives

Table 8. Regression Estimates for Determinants of Site-hitting Behavior[1]

Variables	Est.Coff	Std.Err	t-value	Pr>|t|
Intercept	0.2757	0.3657	0.75	0.4511
CGPA	-0.0033	0.0529	-0.06	0.9511
Credits	0.0017	0.0106	0.16	0.8762
HrsWorked	0.0011	0.0022	0.52	0.6006
ClsParticip	-0.0057	0.0067	-0.85	0.3949
Gender/Male	0.1146†	0.0662	1.73	0.0838
USMinority	-0.2393*	0.1058	-2.26	0.0239
USWhite	-0.2772***	0.0956	-2.90	0.0038
Sophm	0.1020	0.2594	0.39	0.6943
Junior	0.1124	0.2464	0.46	0.6483
Senior	0.0456	0.2472	0.19	0.8532
LSI20	-0.0046	0.0050	-0.92	0.3563
LSI31	-0.0004	0.0050	-0.08	0.9358
HitsOnline$_{i,k-1}$	0.2069***	0.0287	7.20	<.0001

1. N=1092, (Pr>F)<0.0001, R^2=0.39

and minorities. Other variables such as *CGPA, Credits, HrsWorked and ClsParticip* do not show much significance in influencing clicking behavior.

Although site hits might indicate online assignment participation levels it might also indicate a lack of solid conceptual-level mastery of the knowledge. Those who engage in intensive answer checking may be attempting to do the homework by clicking and guessing, rather than understanding the intrinsic knowledge embedded in the course materials.

CHAPTER 5 SUMMARY AND CONCLUSIONS

This work's contribution is the analysis of detailed online behavior data. Various explanatory variables were used, some of which have not been used in previous studies, such as learning styles, site-hit history, and time-management variables. The concept of "marginal benefit" in online learning was introduced to evaluate the specific contribution of each assignment to its associated learning outcome.

We assess the distribution of online learning benefits attributable to different variables – students' characteristics, life style, learning style, efforts expended, and academic proficiency. Among the factors considered, we found significant and positive correlation between the learning outcome and students' efforts in learning the course materials. We are also able to evaluate and identify the marginal benefit for specific assignments. The time spent on traditional learning and students' cumulative school performances are also significant factors in generating learning. Students with higher GPAs, working students, and certain demographic groups benefit more from online learning. No significant differences with respect to course load and academic classification were discovered.

We also investigated the relationship between exam question correctness and topic-related online assignments. We found that topic-related online assignment performance, along with students' general academic performance, significantly contributes to a student's ability to correctly respond to exam questions.

The analysis of online learning participation shows a significant impact from participation in previous assignments, and abstract conceptualization learning preferences. Males and non-native US students tend to seek confirmation through answer checking while doing online homework.

The limitation of this work is that the timespan of the data covers a single semester. Additional data could be included given the recorded data set. Second, this work has collected only certain variables of interest at the time the web site and the evaluation survey were designed. With the completion of this study, other variables of interest have been recognized, e.g. students' familiarity with computers. This weakness can be alleviated by collecting these data of interest now that their importance has been identified.

We draw several implications from this study. First, doing online assignments and class participation, whether in real world or virtual world, needs to be emphasized, as they directly contribute to the final learning outcomes. Grading policies should be designed to encourage participation in online practice problems. Our finding that classroom participation is associated with learning outcomes means that class participation should be rewarded.

One issue of particular interest is that in developing online courses, there is a tendency to develop toward a complete course without questioning the usage and effectiveness of the materials developed. However, no one has analytically sought the combination of traditional and online inputs that optimize the learning objectives. That is to say, we are interested in the percentage of innovative methods such as web-facilitated, multimedia methods in the mix of all learning materials that will lead to the best learning outcomes. It would be interesting to evaluate the variation in online content of different types of online courses. If the lack data on purely online courses could be overcome, then we believe that this would be a good topic for future exploration.

REFERENCES

Allen, E. and Seaman, Jeff.(2010). “Class Differences: Online Education in the United States.” Eighth Annual Sloan Survey of Online Education Shows Economy Still Driving Growth. Sloan Foundation.

Brown, B. W. and Liedholm, C. E. (2002). “Can Web Courses Replace the Classroom in Principles of Microeconomics?” *The American Economic Review*. Vol. 92, No. 2, 444-448.

Dutton, J., Dutton, M. and Perry, J. (2001). “Do Online Students Perform as Well as Lecture Students?” *Journal of Engineering Education*, Vol. 90, Issue 1, 131-136.

Means, B., Toyama,Y., Murphy. R., Bakia, M., and Jones, K. (2009). “Evaluation of Evidence-based Practices in Online Learning: A Meta-analysis and Review of Online-learning Studies.” Washington, D.C.: U.S. Department of Education.

Oblinger, Diana G. (2000). "The Nature and Purpose of Distance Education." The Technology Source (Michigan: Michigan Virtual University) (March/April).

Sonnenwald, D. H., & Li, B. (2003).” Scientific collaboratories in higher education: Exploring learning style preferences and perceptions of technology.” *British Journal of Educational Technology*, 34(4), 419–431.

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.

Taraban, R., Maki, W. S., & Rynearson, K. (1999). “Measuring study time distributions: Implications for designing computer-based courses.” *Behavior Research Methods, Instruments, & Computers*, 31, 263–269.

Young, S. S. (2004). “In search of online pedagogical models: Investigating a paradigm change in teaching through the School for All community.” *Journal of Computer Assisted Learning*, 20, 133–150.

APPENDIX A VARIABLE SUMMARY

Table A.1. Number of Observations by Assignment

Variable	N
Online Simulation Assignment #0	75
Online Simulation Assignment #1	81
Online Simulation Assignment #2	76
Online Simulation Assignment #3	70
Online Simulation Assignment #4	66
Online Simulation Assignment #5	60
Online Simulation Assignment #6	65
Online Regular Problem Set #2	83
Online Regular Problem Set #3	75
Online Regular Problem Set #4	75
Online Regular Problem Set #5	59
Online Regular Problem Set #6	51
Online Regular Problem Set #7	55
Online Regular Problem Set #8	64

APPENDIX B SAMPLE COMPOSITION

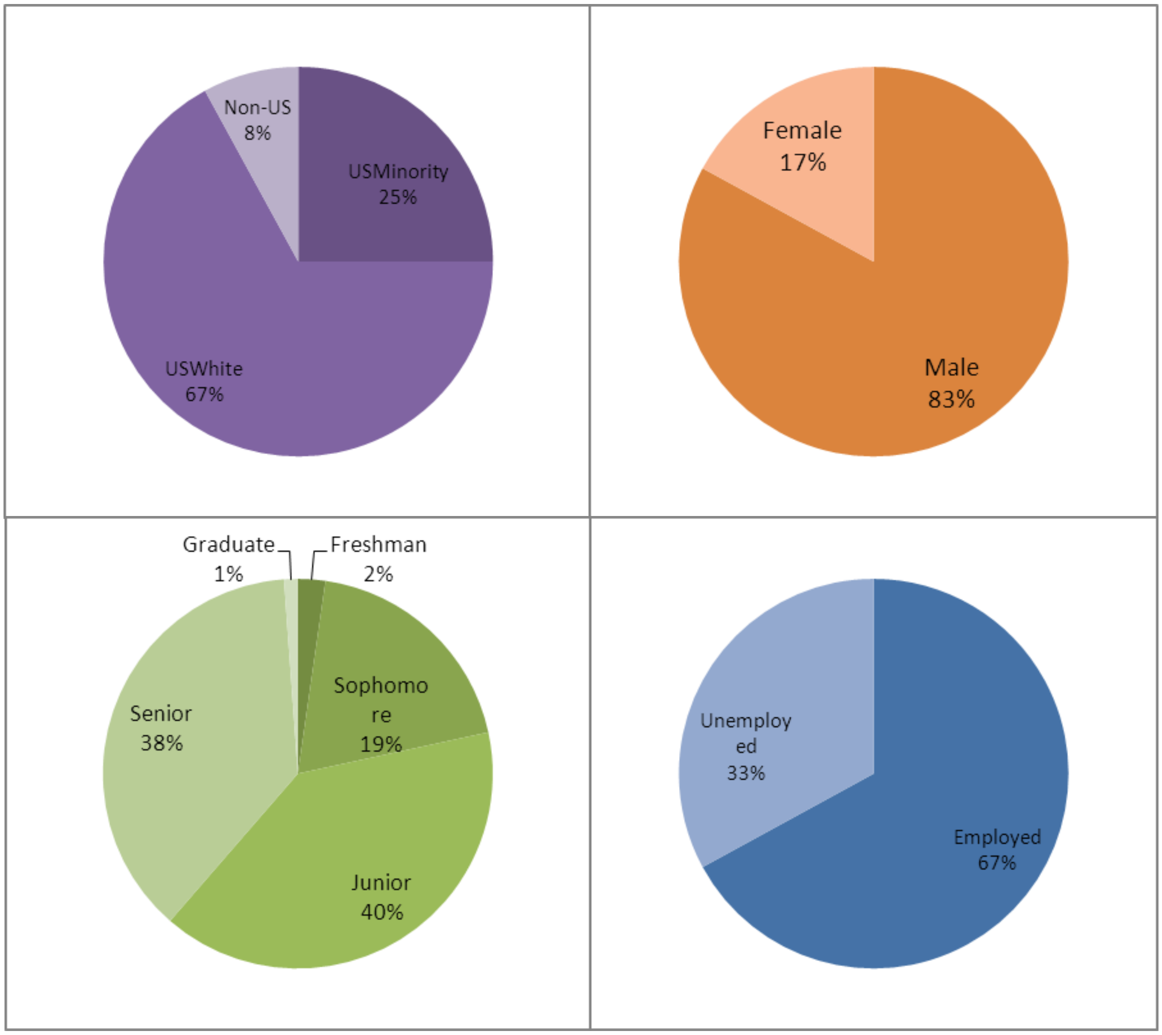

Figure B.1. Pie Charts of Data of Sample Composition

APPENDIX C HOMEWORK-EXAMINATION ITEM PAIRINGS

Table C.1. Homework and Exam Item Matching Pairs

	MT1	MT2	Final
1			
2			
3			
4			
5			
6			HW3, parts 1, 4
7	HW2, all parts		
8	HW2, all parts		
9			
10			
11		HW5, given P find disc yld	HW4. FTE1. FTE1.
12		HW5, 2 yr T bond, find price	
13			
14			
15			Form b: a. HW5, parts 8 and 9. b. HW5, given d find P. c. HW5, given d find P. d. HW5, parts 8 and 9. e. HW5, parts 8 and 9.
16		HW6, parts 2, 3, 4	HW6, part 1
17	HW3,part2 and part 5		
18			
19			HW6, part 7. FTE4
20		HW6, part 6	
21		FTE4	HW7, part 5. FTE5
22			
23	HW4, all parts	HW7, part 1	
24		HW7, part 5. FTE5	HW8
25			
26			FTE6.
27			FTE6
28			
29			
30			

APPENDIX D EXAM ITEM CORRECTNESS STATISTICS

Table D.1. Exam Item Correctness Statistics

Pair	Mean	StdEv
MT1HW2	0.5000	0.5030
MT1HW2_2	0.4288	0.4978
MT1HW3	0.7976	0.4042
MT1HW3_2	0.7143	0.4545
MT1HW4	0.7500	0.4356
MT2HW5	0.8929	0.3112
MT2HW5_2	0.5119	0.5029
MT2HW6	0.5952	0.4938
MT2HW6_2	0.2262	0.4209
MT2HW7	0.7143	0.4545
MT2HW7_2	0.4048	0.4938
MT2FT4	0.3452	0.4783
MT2FT5	0.4048	0.4938
FINHW3	0.4167	0.4960
FINHW4	0.5119	0.5029
FINHW5	0.4048	0.4938
FINHW6	0.1548	0.3639
FINHW6_2	0.3810	0.4885
FINHW7	0.4167	0.4960
FINHW8	0.7024	0.4600
FINFT1	0.5119	0.5029
FINFT4	0.3810	0.4885
FINFT5	0.4167	0.4960
FINFT6	0.5238	0.5024
FINFT6_2	0.3333	0.4742