Fall 11-15-2017

Which factors influence student success in Intermediate Algebra, MATH 101-102-103?

Linh T. Ward
University of New Mexico

Linh Ward
Candidate

Mathematics and Statistics
Department

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Erik Erhardt, Chairperson

Terry Loring

Cristina Pereyra

Which factors influence student success
in Intermediate Algebra, MATH 101-102-103?

BY

LINH WARD

BS, Mathematics, The University of New Mexico

THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Master of Science

Mathematics

The University of New Mexico
Albuquerque, New Mexico

December, 2017

DEDICATION

For my husband, Jerry, who went to Vietnam for me, brought me to America, and has supported and encouraged me to go to school.

ACKNOWLEDGMENTS

First, I would like to thank you Professor Erik Erhardt, for kindly being my advisor. Professor Erhardt spent a lot of time helping me, through the birth of his son, labored over and helped me with my writing, and taught me many of the skills needed and necessary to apply for my thesis.

I would like to thank Professor Maria Cristina Pereyra and Professor Terry Loring who introduced me to Professor Erhardt.

Thank you Professor Maria Cristina Pereyra for all the support you gave me along with all the students in the Mathematics and Statistics department.

I want to thank Doctor Srini Vasan who was always there for me with answers to my questions about ALEKS and questions about MATH 101 – 102 – 103 for my thesis.

I want to thank to Ana Parra Lombard, the special adviser of Mathematics and Statistics Department for all of the support you gave me along with all the student in the Department.

I want to thank Sonia Gipson Rankin for helping with my questions, always asking if I needed help or information, and encouraging me and helping with my thesis.

Thanks to Eric Ehrhardt Hamke , a very good friend who always help whenever I need for the help with my thesis.

All these wonderful and dedicated people helped me immensely with my thesis and turned the dreaded trials and tribulations into a learning experience. Without the help, dedication, and commitment these people made to help, there would not be this dissertation to be presented here.

There are so many others that provided the proverbial pat on the back; the people who said you can do it encouragement, others that helped with the language hurdles and confusion, and just so many more that just wanted me to succeed. Thank you all.

Which factors influence student success

in Intermediate Algebra, MATH 101-102-103?

by

Linh Ward

BS, Mathematics, The University of New Mexico

Abstract

At The University of New Mexico (UNM), Intermediate Algebra (MATH 120 and MATH 101-102-103) has historically been a so-called “killer course”, with very low pass rates: approximately 40% in Fall 2009 to Spring 2011 and about 50% from Fall 2011 to Spring 2013. Furthermore, many students failed the class multiple times. Since 2013, a computer system called ALEKS has been used to teach the course and, along with some additional interventions, on Albuquerque/Main campus success rates for MATH 101 have increased to roughly 80% and MATH 102 to about 70%.

This thesis provides a strategy to identify those 20-30% as-risk students most likely to need additional support to succeed. By combining data from the UNM Registrar (Grades, ACT scores, and demographics), NM county-level poverty data, and response-level ALEKS assessment and practice metrics, we developed a statistical model that uses data from the first week of class and

predicts with almost certainty (1% overall error) whether a student will pass in that semester. This represents another potentially important incremental improvement to a series of successes in redesigning Intermediate Algebra at UNM.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 Introduction

At The University of New Mexico (UNM), Intermediate Algebra (MATH 120 and MATH 101-102-103) has historically been a so-called “killer course”, with very low pass rates: approximately 40% in Fall 2009 to Spring 2011 and about 50% from Fall 2011 to Spring 2013 (Figure 3). Furthermore, many students failed the class multiple times. Since 2013, a computer system called ALEKS has been used to teach the course and, along with some additional interventions, on Albuquerque/Main campus success rates for MATH 101 have increased to roughly 80% and MATH 102 to about 70% (Figure 3). Given that we still aim to help every student succeed, this thesis answers this question: Which factors influence student success in Intermediate Algebra, MATH 101-102-103? The goal is to develop a statistical model to help predict which students are most likely to need additional support so that early intervention may improve their success, and therefore further improve the overall success rate in the course.

We make three main points in this thesis. First, by adopting the computer-based system ALEKS and converting MATH 120 to MATH 101-102-103, there was a substantial increase in the student success rate and decrease in students repeating the course. Second, course coordinator Dr. Srini Vasan and others have made improvements to the course each semester for additional incremental increases in student success. Third, the statistical model developed in this thesis can be used on the first day of class to predict whether a student will pass with a 1% error rate to identify those students most likely to fail for the purposes of intervening to help those at-risk students succeed. This represents another potentially important incremental improvement to a series of successes in redesigning Intermediate Algebra at UNM.

1.1 UNM Students

UNM is a Hispanic-serving institution with many low-income students (The University of New Mexico, 2017). Note that in Spring 2017, 7 of the top 15 highest failing-rate courses at UNM are Mathematics courses (MATH 121, 150, 162, 153, 102, 180, and 123) all with failure rates over 40% (The University of New Mexico, 2017) Mathematics placement at UNM.

Before Fall 2012, Intermediate Algebra (MATH 120) was the highest level mathematics course before taking core math or statistics. Each year, approximately 2500 students enrolled in Intermediate Algebra. This course was the prerequisite for all mathematics courses satisfying university-level general education requirements (Wang, 2014). This course did not satisfy UNM core requirements but was used as an elective preparatory course (The University of New Mexico, 2017). Each semester a substantial proportion of new students seeking to enroll at UNM do not meet general education requirements and need rudimentary preparation; 25% of first-time freshman in Fall 2015 needed Foundational Math (UNIV 103) (Rankin, 2016) and Figure 8 suggests that a larger proportion start in MATH 101.

Mathematics placement is based on American College Testing (ACT) scores. Before Fall 2015, students with a Math ACT score of 11-18 were placed into remedial Introductory Studies Math (ISM), 19-21 into MATH 120 (MATH 101-102-103) which are below core math courses, and 22+ into a core math course; since Fall 2015 a Math ACT score of 18 now places students into the higher MATH 120 (Figure 1) (College, 2017). Historic pass rates for the traditional lecture-based Intermediate Algebra have been around 45% (Figure 3). Before Spring 2013, UNM had approximately 20 sections of Intermediate Algebra, each with about 60 students. Based on Wang's thesis, "Instructors had tried many different approaches and strategies to improve student performance, including a completely internet-based section with online homework sets using commercial software, and integrating individual or group work into the lecture via handouts or a workbook. Since no formal evaluations of these alternative methods of instruction have taken place, no attempted changes have been widely implemented" (Wang, 2014).

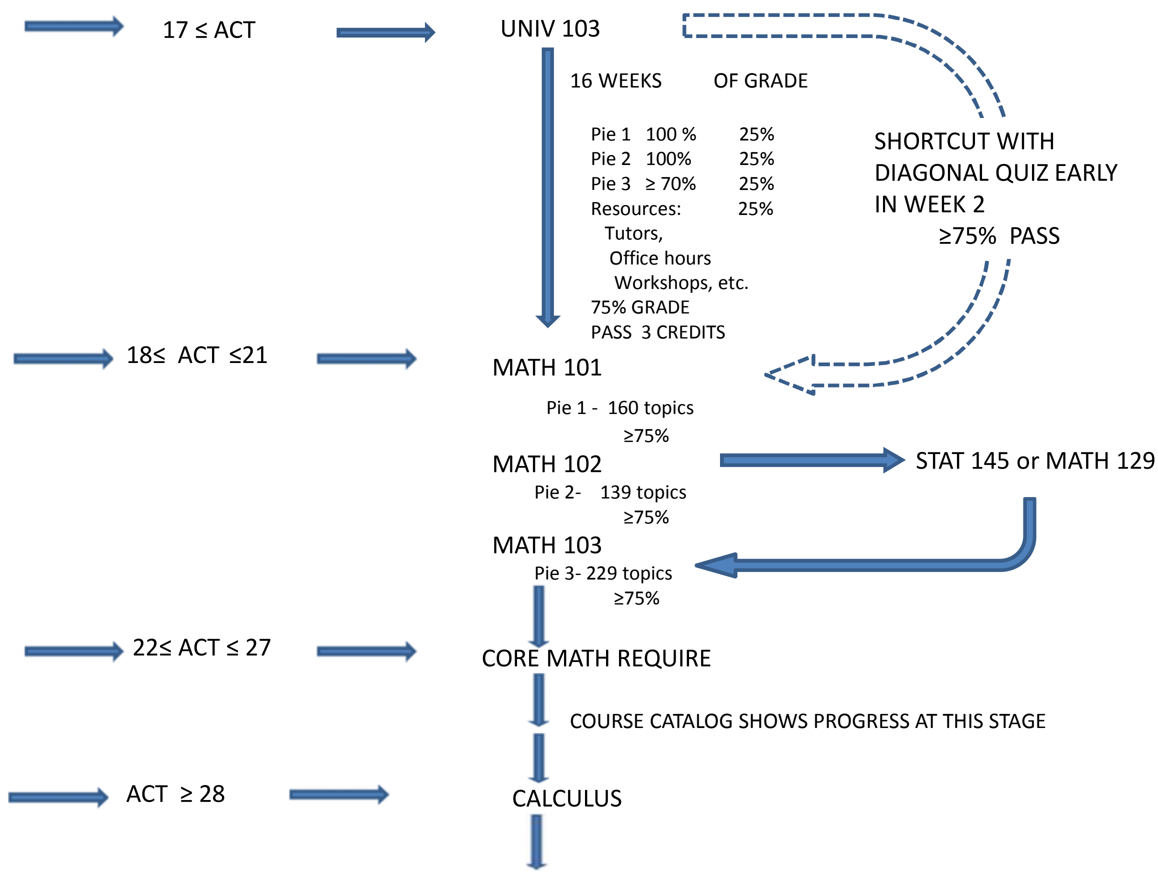

Figure 1 Pathway to mathematics courses based on incoming ACT scores and course requirements since Fall 2015.

1.2 *Historical Success Rates*

Albuquerque's main campus is by far the largest campus providing MATH 101 instruction, and of the six largest campuses, Gallup is the only campus to continue teaching MATH 120 instead of transitioning to MATH 101, and Valencia campus teaches both (Figure 2). Below, we define "success" as receiving a letter grade of A+ through C and CR (credit), "failure" as a C- through F and I (incomplete), NC (no credit), W (withdrawl), and NR (no progress), and "neutral" as PR (progress) and AUD (audit). We exclude "neutral" and calculate the proportion of students who succeed. Figure 3 illustrates that on Albuquerque /Main Campus MATH 101 has the a highest success rate, MATH 102 has the lowest success rate (these are students who previously passed MATH 101) and that the success rate of MATH 102 is not that different from the historical success rate of MATH 120.. Note that the success rate at Gallup with MATH 120 is as high as Albuquerque with MATH 101. Figure 4 illustrates that the mean GPA also increased by nearly a full letter grade from MATH 120 to MATH 101 in Albuquerque, though with more modest increases at other campuses. Finally, Figure 5 illustrates the proportions of whole letter grades excluding students who did not complete the course.

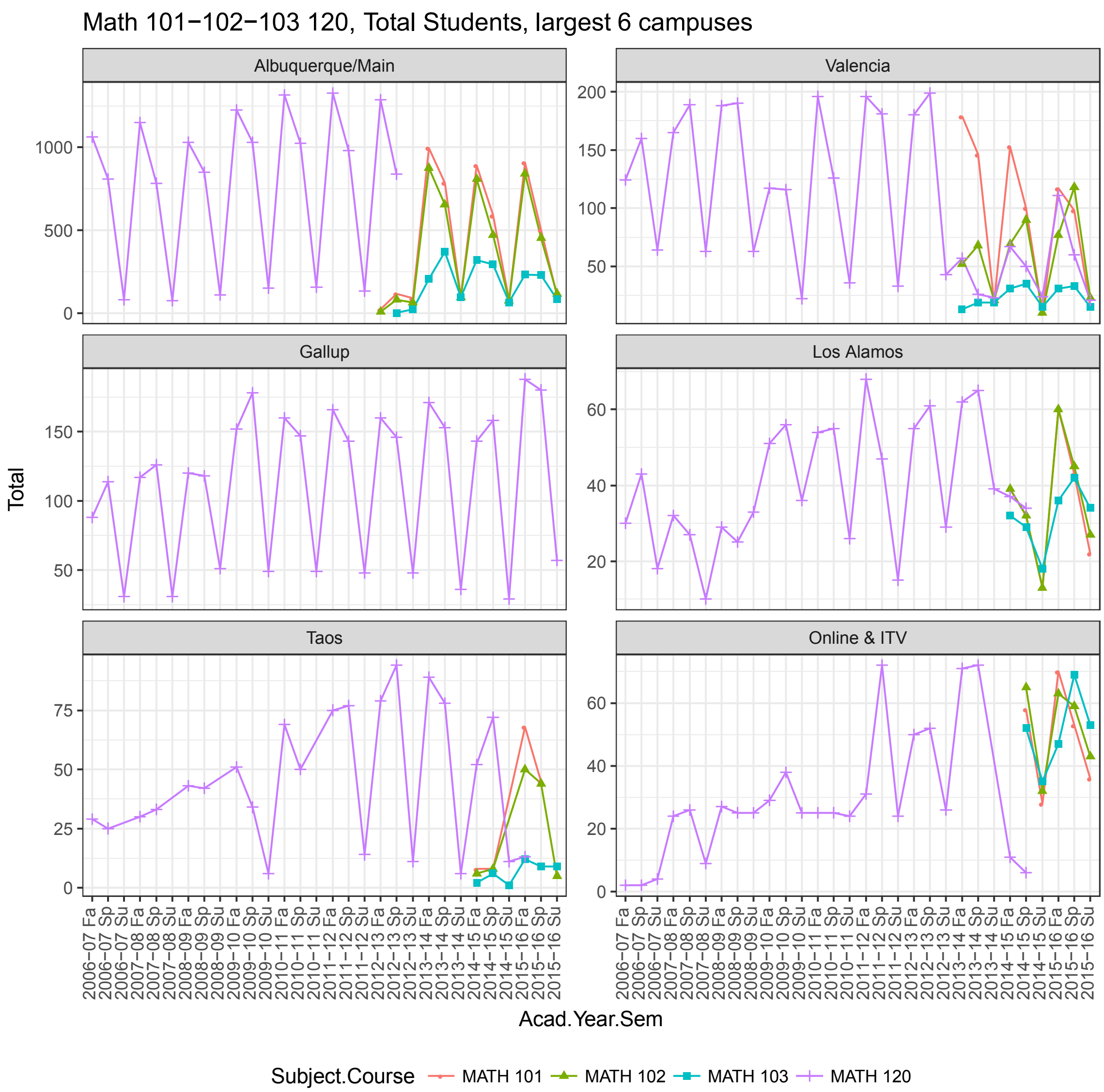

Figure 2 Historic total number of students taking MATH 120 or 101-102-103 for the six largest UNM campuses, common y-axis.

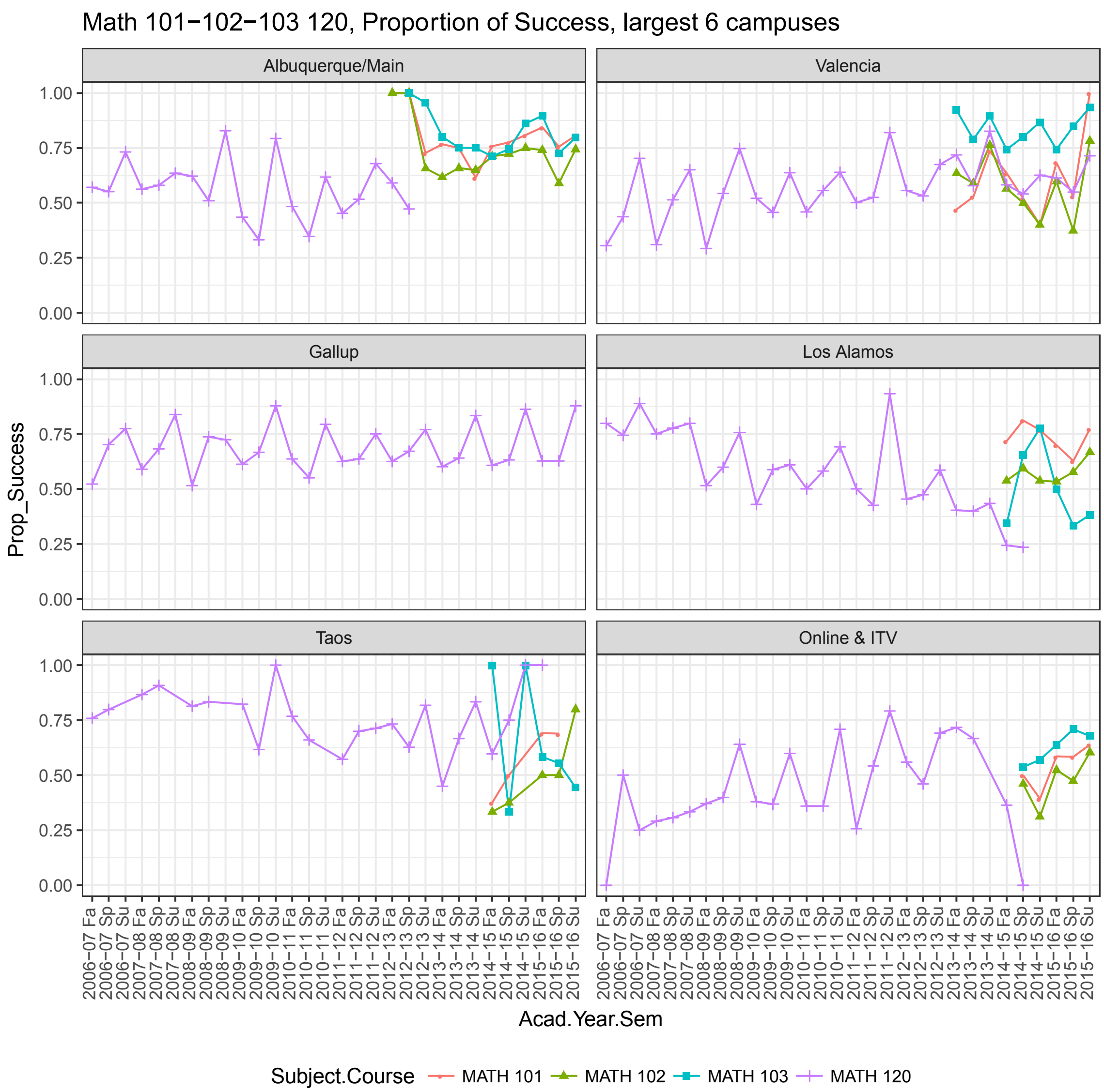

Figure 3 Historic of the proportion student who success of taking MATH 120 or 101-102-103 for the six largest UNM campuses.

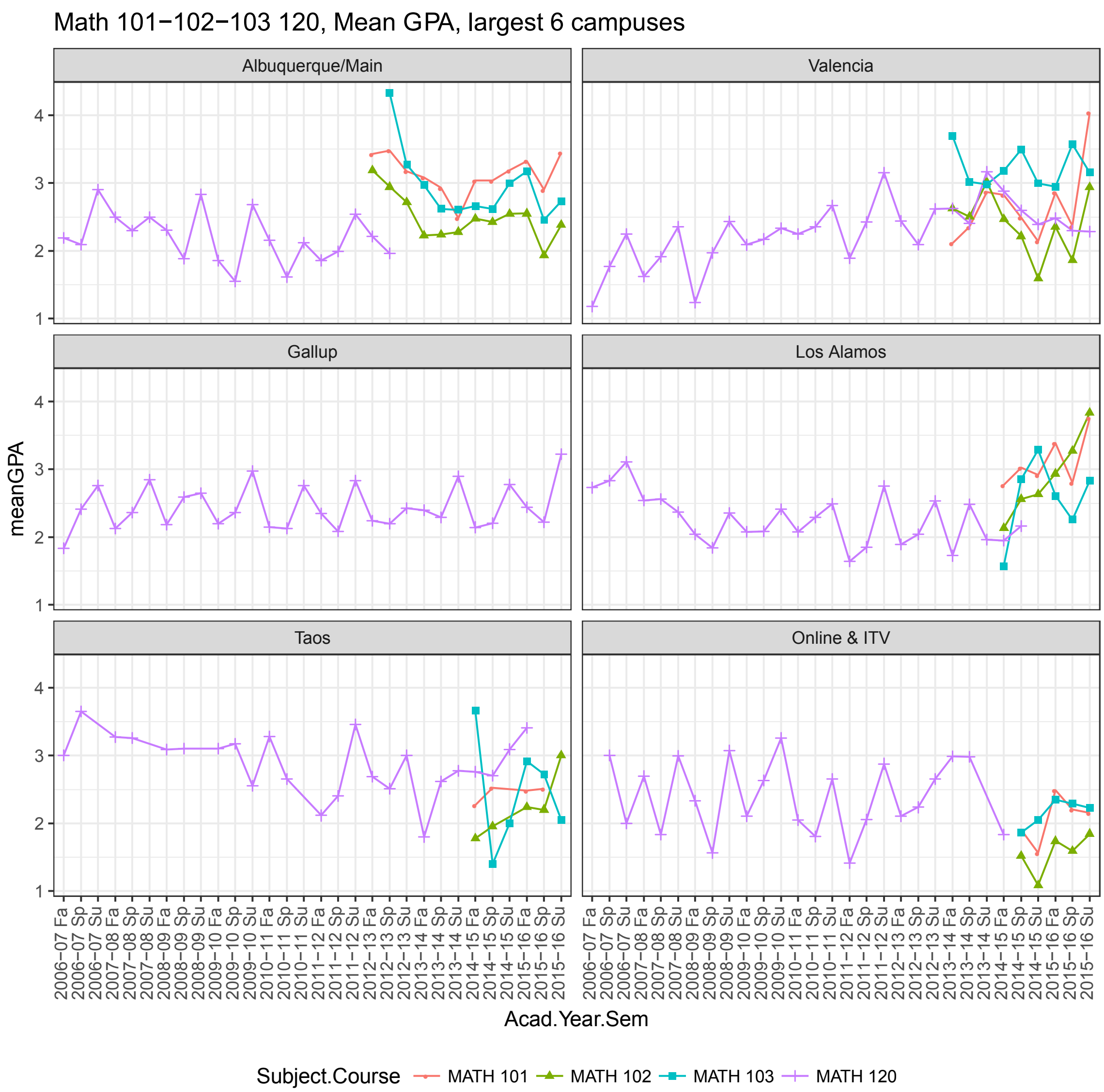

Figure 4 Historic of the mean GPA of students taking MATH 120 or 101-102-103 for the six largest UNM campuses, common y-axis.

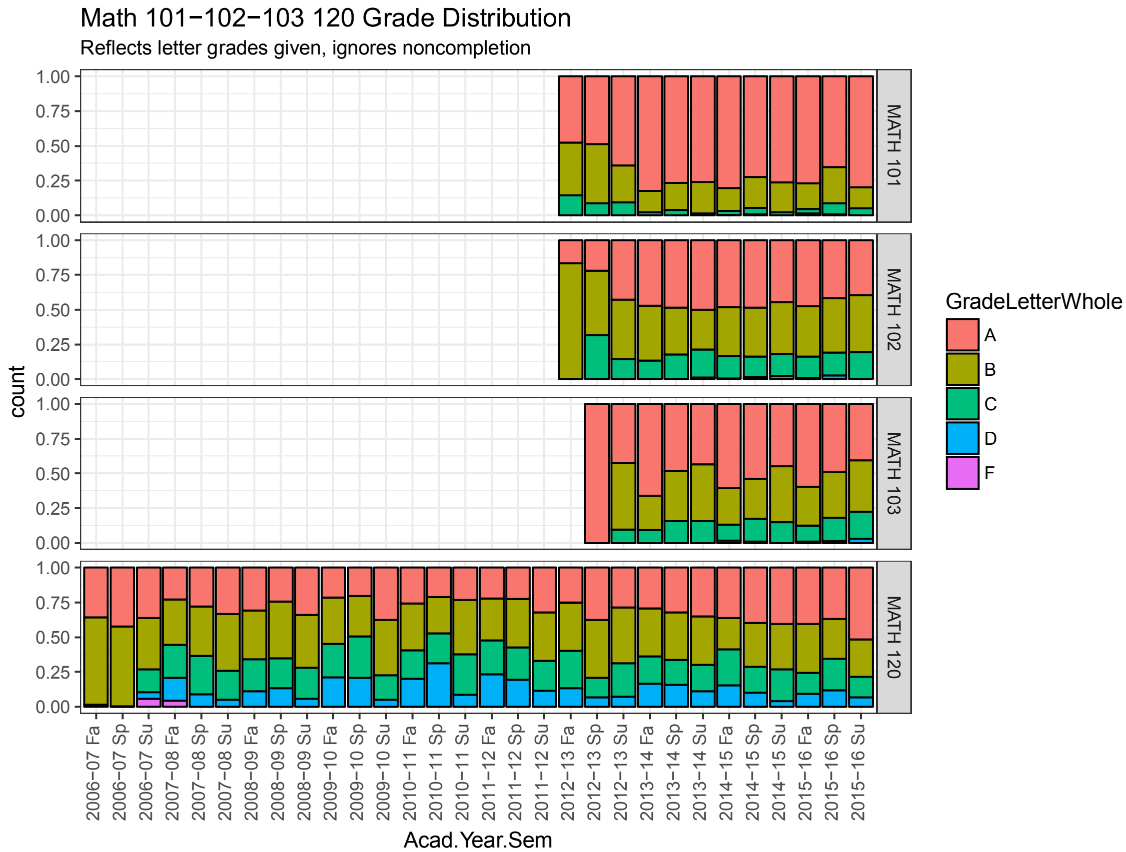

Figure 5 Historic grade distribution of students taking MATH 120 or 101-102-103 for the six largest UNM campuses

1.3 The development of a modern Intermediate Algebra course

In Spring 2012, a team of University administrators, faculty, and graduate students researched different instructional approaches used at other institutions, with the intent of increasing the success or pass rates of students in both Intermediate Algebra and subsequent required courses (Wang, 2014). The team determined that there was significant promise in one particular computer-based model and proposed to pilot it in Fall 2012 (Wang, 2014). In spite of the pilot not being completely successful, the university administrators accelerated the full implementation of the model in Spring 2013 (Wang, 2014).

In Spring 2013, no traditional version of Intermediate Algebra MATH 120 was offered at UNM Main/Albuquerque campus. Instead, the redesigned self-paced course used the web-based computer software ALEKS (discussed below) and was offered in a repurposed space in called the Math Mαll Lab located in the Centennial Library (L185) (Steve Carr, 2013). The Math Mαll Lab was equipped with 140 computers divided into two sections: the 125-computer classroom for practice problems and the 15-computer testing room for assessments (Steve Carr, 2013). There are also open offices for the course coordinator, instructors, and student tutors who work at the Math Mαll Lab. The pass rate for that first year was still low, and many students still could not complete the course (Vasan, 2016); specifically, the last MATH 120 point in Spring 2013 in Figure 3 is in the Math Mαll Lab.

In Summer 2013, the 3-credit 16-week (1-semester) Intermediate Algebra (MATH 120) course was split into the series of 1-credit 8-week MATH 101-102-103 individual courses. Each semester at UNM's main campus is split into two 8-week terms. Commonly, MATH 101 is offered during the first 8-week term in a semester and MATH 102 (with optional MATH 103) during the second 8-week term. During the first 8-week term of the semester, there are about 9 sections of MATH 101 and about 2 sections of combined MATH 102-103. The second 8-week term is the opposite with about 2 sections of MATH 101 and about 9 sections of combined MATH 102-103. Each section of MATH 101 has about 100 students, while MATH 102-103 may have fewer. There are some restrictions to register for MATH 102-103 in the registration. The content of Intermediate Algebra was separated in this way:

- MATH 101: Linear equations and inequalities, applications and problem solving with linear equations, linear functions and the graph of a line, percent, perimeters, and areas of simple geometric shapes.
- MATH 102: Quadratic equations, properties of exponents and scientific notation, simplifying polynomial expressions, factoring, and introduction to functions.
- MATH 103: Radical expressions and equations, rational expressions and equations, and the exponential and logarithm functions.

There were good reasons for breaking MATH 120 into MATH 101-102-103. There was more content in MATH 120 than many students needed in order to take either of the core classes Introduction to Statistics (STAT 145) or Survey of Mathematics (MATH 129) (The University of New Mexico, 2017) now MATH 101 and 102 were sufficient. Furthermore, having separate 1-credit courses allowed students more than one semester to complete the full 3-course 101-102-103 sequence, benefitting students who needed more time to succeed (Figure 3).

A big initial challenge was the burden of handling incomplete (I) grades at the end of the semester (Vasan, 2016c). Students were allowed 4 months after the end of a course's term to complete each course in the series (Ross, 2014). It became a strain on the course coordinator at the end of each semester because the instructor contracts terminated at the end of the semester so handling all the student grade changes became the responsibility of the course coordinator (Vasan, 2016c). All incomplete grades needed to be changed either to a letter grade if the student completed the course successfully or to no credit (NC) after 4 months. In Spring 2016, students were no longer allowed to take an incomplete grade at the end of each course, but were given the additional time of the break interval between semesters to earn a letter grade to avoid a no-credit (NC) grade and needing to reregister to take the course again (Ross, 2013).

While the content has remained stable, many support resources have been developed to meet students' needs (Vasan, 2016c) which has resulted in higher pass rates and higher grade point averages (Figure 3, Figure 4, and Figure 28), and a decrease in the cost of instruction per student since some students don't need the third credit of MATH 103 and many fewer students need to retake the course.

There are up to four types of helpers in every classroom (Vasan, 2016b). The course coordinator and the instructors are in charge and available in the lab to provide help to students if they are struggling. Instructors are also available to help students with registration, content, etc. Tutors are student helpers who are also available to help students one-on-one at a ratio of about 1 tutor for every 15 students. For students who really struggle, a student peer mentor may be assigned to them to provide support and individual help.

Because students are given multiple attempt on exams they do poorly on, rather than simply failing at the end of the semester (as in MATH 120), many will succeed after an additional attempt or after more practice (Vasan, 2016c). The Math Mαll Lab is open Monday through Friday from 8 AM to 6 PM and Sunday from 12 PM to 6 PM with 15 computers in the testing room. A student can take an exam any day when the Lab is open provided there is enough time to complete the first exam attempt (Vasan, 2016b). A student has 2 attempts for an exam after completing a set of topics (which ALEKS calls a “Pie”). The exam does not have a time limit. After taking their first exam attempt and getting their grade, a student has a chance to look back over the exam they completed. That student then has time to ask questions and study on their own before returning and retaking the exam questions that they missed. The highest score for each question will be counted toward the final grade for each exam. If after two attempts, that student still does not score 75% or higher, the instructor will request the comprehensive knowledge check (assessment) to help the student relearn the topics and complete their “pie” again. The student can still look over the exam they did and study. Finally, the student returns and has a fresh exam with two attempts available to them (Ross, 2013).

The Math Mαll's website includes course resources, such as testing review, workshop schedules, instructional videos, Math Mαll hours and location, and other Math Department and UNM resources (The University of New Mexico, 2016). A Math Mαll Facebook page was created in the Fall 2016 with YouTube video lectures for some difficulty topics. Furthermore, there is no textbook requirement for the course because all materials are online in the ALEKS software (Ross, 2013).

1.4 ALEKS:

ALEKS (**A**ssessment and **LE**arning in **K**nowledge **S**paces) is a Web-based, artificially intelligent assessment and learning system used primarily for mathematics, but also statistics, accounting, and chemistry (McGraw Hill Education, 2017c). ALEKS was developed by using Knowledge Space Theory (Falmagne, Cosyn, Doignon, & Thiéry, 2006). A group of software engineers, mathematicians, and cognitive scientists at New York University and the University of California, Irvine, developed ALEKS under several National Science Foundation (NSF) grants which track what a student knows from their previous responses to questions and predicts which questions they are ready for (Figure 6). This allows a course to be changed from a traditional lecture to an adaptive computer web-page. Many courses have been taught using ALEKS, such as mathematics, science, and business, from K-12 through college level. Since it is a Web-based, millions of people use ALEKS for learning all over the world (McGraw Hill Education, 2017a). “The Average Historical Student Learning Rates with ALEKS are ~90%” (McGraw Hill Education, 2017a).

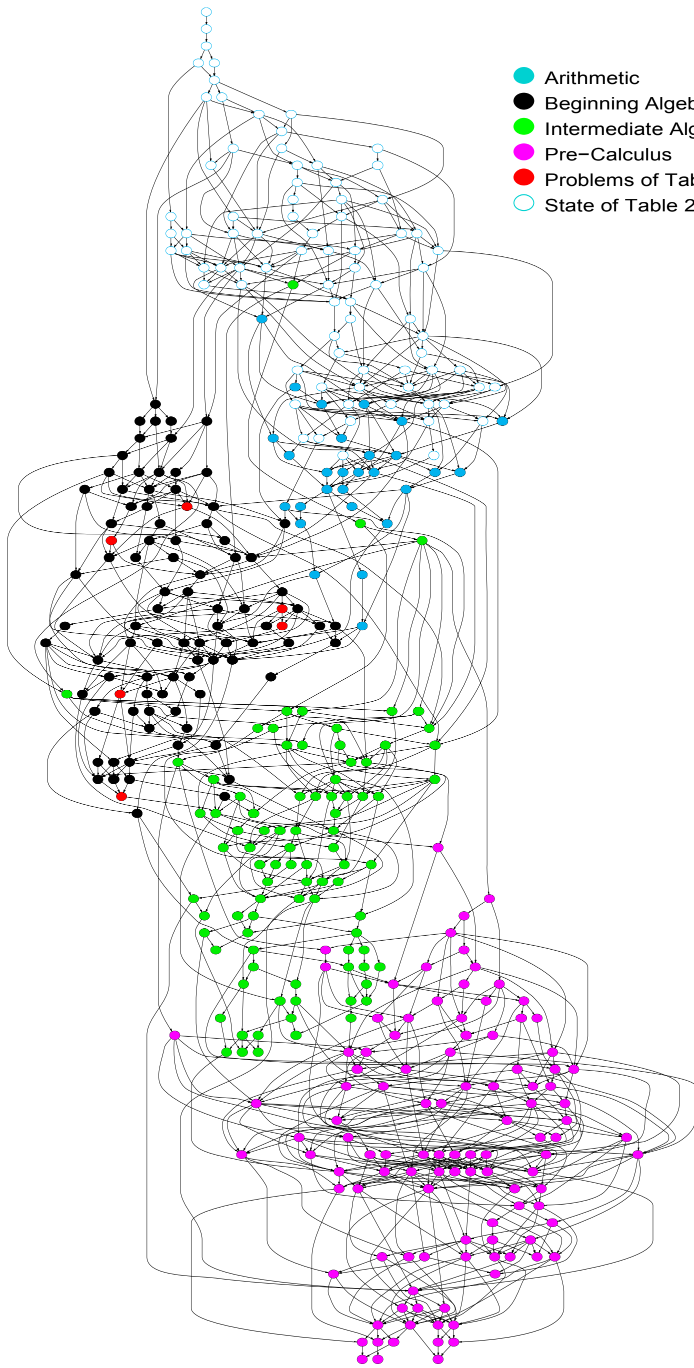

Figure 6 Each question in ALEKS requires the skills evaluated from previous questions (taken from (Falmagne et al., 2006)

ALEKS has been used successfully at many schools (McGraw Hill Education, 2017b). After moving from a lecture-based format to ALEKS' online adaptive format, both high school and college math instructors have seen pass rates improve (roughly double), have seen skill proficiency jumps of a few grades over a single semester, and have seen increased student retention (McGraw Hill Education, 2017b). This program has made a positive impact on students learning, adapting to a specific student's weaknesses and bringing their skills from bad or weak in algebra to being able to understand the concepts and feel confident to learn math (McGraw Hill Education, 2017b). Questions are presented in either of two languages, English and Spanish (McGraw Hill Education, 2017b).

At UNM, after creating an account in ALEKS, a student will get an initial 20-30 question assessment to test how much the student knows in the course (Ward, 2017). The assessment adapts so that each subsequent question depends on previous answers (Ward, 2017). The more questions the student answers correctly, the more questions ALEKS will ask (Ward, 2017). Then, ALEKS will analyze the student's knowledge and summarize the results with a pie chart presenting what the student knows and needs to learn in the course (Figure 7). The dark color of a pie chart represents what the student knows, and the light color represents what the student does not know and needs to learn. The pie chart for each student will be different depending on each student's knowledge of the material that ALEKS assessed. "As a student works through a course, ALEKS randomly reassesses the student, and then the pie chart will be adjusted depending on the result of the assessment" (McGraw Hill Education, 2017c). ALEKS includes a set of open-ended questions, so that students need to work through a problem to give a typed-out answer, rather than selecting from multiple choices. Open-ended questions are more effective at assessing a student's understanding, and the algorithms behind ALEKS help to present material in the right increments of difficulty to prepare them for the next level (Falmagne et al., 2006).

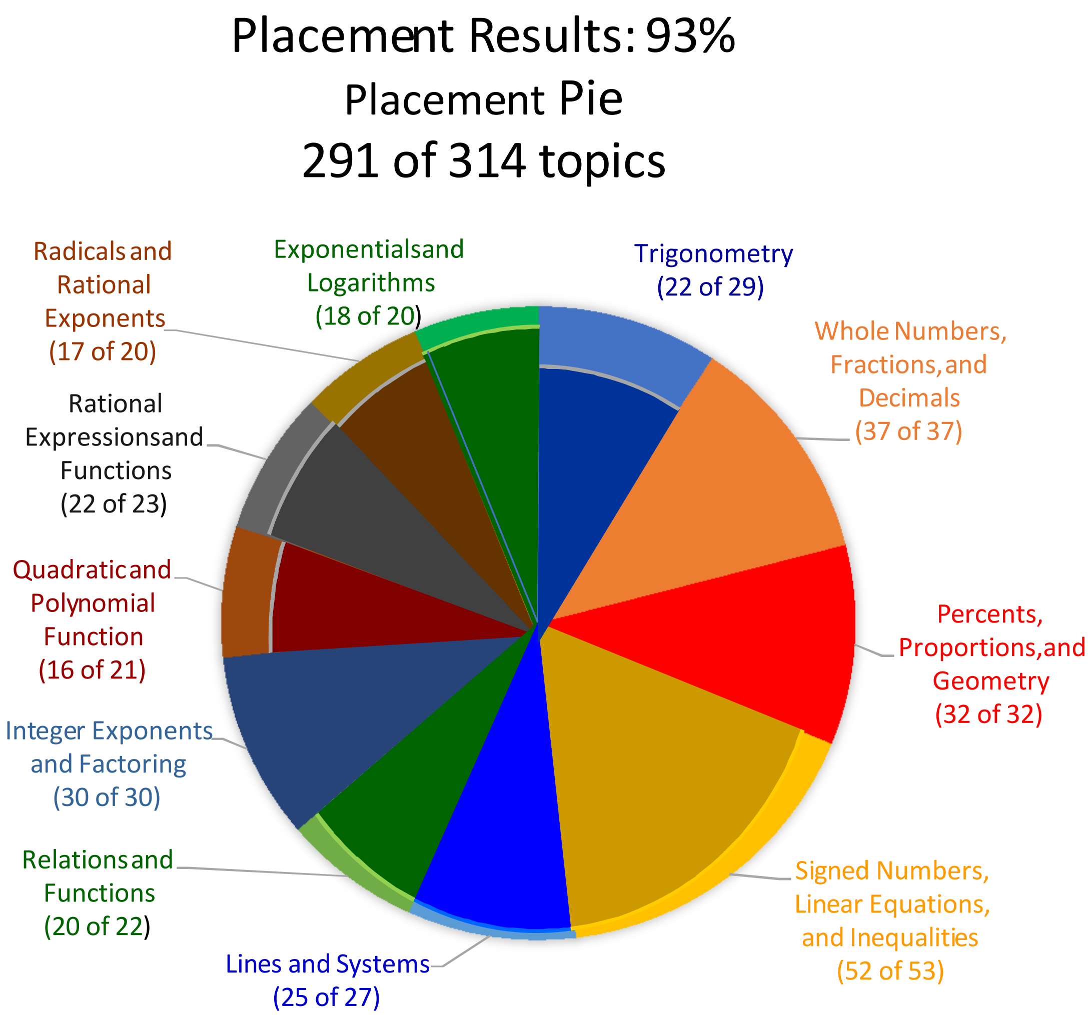

Figure 7 ALEKS Pie Chart of topics and attained mastery

1.5 Larger impacts of ALEKS at UNM

UNM uses ALEKS in a few courses. UNIV 103 (Foundational Math, previously UNIV 102 Quantitative Reasoning) adopted ALEKS in Fall 2015 for the pre-college preparatory levels of mathematics offered at UNM. MATH 101-102-103 (Intermediate Algebra) started using ALEKS in Fall 2012 for the college preparatory levels of mathematics. An introductory chemistry course also uses ALEKS. The math courses will be discussed here.

[redacted] *UNIV 103 (Foundational Math)*

Annually, over 500 students enter UNM underprepared for introductory mathematics courses. UNIV 103 concentrates on 4 goals (Rankin, 2016):

- Goal 1: Refine placement into Foundational Math
- Goal 2: Reduce time to starting core math course
- Goal 3: Prepare students for the Math and Learning Lab (MαLL), MATH 101-102-103
- Goal 4: Introduce students to academic foundational success skills

For Goal 1, since Fall 2015, an ACT Math score of 17 or below would place a student into UNIV 103 (Foundational Math) instead of ACT Math score of 18. Also, students can bypass UNIV 103 if they score of 75% or higher on the “diagonal quiz” of UNIV 103, a quiz that students take in week two of each semester to assess whether they should be placed higher into MATH 101 or not.

For Goal 2, department advisors ensure that each student takes a math course in their freshmen year to be on track for completing all their course requirements on time for their degree. Historically, some students would defer math requirements until their senior year and then fail and be stuck at the end of their degree.

To help meet Goal 3, in Fall 2015, University College took over “Introductory Studies – Mathematics” (ISM 100) from a special remedial unit and redefined the curriculum using ALEKS (Rankin, 2016). The design was similar to MATH 101-102-103 which helped students become familiar with the computer-based class format. UNIV 103 is a 3-credit course and contains three sets of topics to complete, called “pies”. Each pie is average 100 math topics which the students should master. Pie 1 and half of Pie 2 is foundational arithmetic to help students in MATH 101. Half of Pie 2 and Pie 3 are preparatory for MATH 102.

For Goal 4, Foundational math utilizes the Success Navigator survey as a reflective tool to identify individual student areas of strength and improvement (i.e., connection to college, self-efficacy, stress management, etc.) (The University of New Mexico, 2017). Students complete a follow-up assignment directly related to their Success Navigator results as an opportunity to build their resources, skills, and/or knowledge in a specific area regarding college success. These academic foundational skills are developed by successfully discovering and using UNM community resources, working in a group, completing each pie on time, attending campus events, scheduling appointments, or using office hours to see tutors and teachers, etc.

The final course grade is 25% for each exam after completing each of the three pies, and 25% for academic foundational skills (Rankin, 2016).

1.6 Literature Review

An important obstacle to student success is the difference between high school and college learning obligations (Knowlton, 2011). Many high school students in their junior or senior years do not have to take math classes; they can take courses they enjoy more or less stringent courses to obtain a better GPA. While this may make high school either more pleasant or easier it does not make college entry requirements either pleasant or easy. This is particularly true for college mathematics courses (Knowlton, 2011). Many colleges have instituted remediation mathematics classes to bridge high school skills to college entry skills (Brown, R. S., & Niemi, 2007). Nearly 20% of entering college mathematics students find themselves in remedial classes (Complete College America, 2012) but many of these attempts are unsuccessful. A study of remedial students in community colleges in Florida found that they were apt to take additional mathematics courses but no more likely to graduate or transfer to a four-year college (Calcagno & Long, 2008). Further, even if students had a pre-calculus course in high school and took a pre-calculus course after entering college they appeared to have at least a 6% reduction in their grade comparing their high school and college grades (Bressoud, 2014) . The remedial classes students take to prepare them to take regular college classes do not count for college credit (Cohen, Brawer, & Kisker, 2014). Based on a report by Complete College America, for students who require remedial courses in math and English, only 17% will graduate (Complete College America, 2012, 2016). Furthermore, applying a strategy called “co-requisite remediation”, where students enroll directly into college-level courses and receive academic support alongside their regular classes, the rate of students enrolled in remedial classes who go on to complete their associated introductory (gateway) course goes from 22% taking 2 years to over 60% taking 1 or 2 semesters (states participating: GA, IN, TN, WV, and CO) (Complete College America, 2012, 2016).

The City University of New York (CUNY) reports that they require 60% of entering students to take remedial math courses in order to be able to pass college-level math courses. Many students delay this requirement and this becomes an impediment when the students are close to graduation and have not taken the remedial program in order enter into the mathematics courses that are required (Bailey, T., Jeong, D. W., & Cho, 2010). Thus, students who enter CUNY and are required to take remedial math courses have a lower graduation rate than students who do not need remedial classes (Attewell, P., Lavin, D., Domina, T., & Levey, 2006; Complete College America, 2012). In California community colleges, over 70% of students are placed in remedial math (Brown, R. S., & Niemi, 2007).

While UNM has discontinued Introductory Studies courses (Suilmann, 2015), for the purposes of this thesis we define “remedial” as any course taken that does not earn the student credit towards graduation (The University of New Mexico, 2017). At UNM the remedial mathematics program includes UNIV 103, MATH 101-102-103, and MATH 120.

In a traditional classroom at UNM, a standard class size has the capacity of sixty seats in Dane Smith Hall, the large classroom building. For entering students, a large size reduces the chances that students feel the need to interact with the instructor or to problem solve, and fewer students have the chance to participate in or discover the experience of learning (Cuseo, 2007). Students in the 21st century may not be interested in a traditional lecture classroom because of the fast development of technology (Cuseo, 2007). To accommodate the high demand of remedial courses, US colleges have been using a new course format for basic college mathematics courses called “emporium” (Carol A Twigg, 2011). According to Twigg, the benefits of emporium classroom are:

- Students spend time doing problems right away in each class instead of watching the teacher solving problems in their lecture,
- Each student starts at a different point and can pass by what they already know, instead spending time developing skills they don’t have yet,
- Students have the opportunity to interact with tutors one-on-one on the problem they struggling with, and
- Students can read the explanation and watch and video (YouTube) as their “tutor” when not in the classroom.

These emporium courses are held in large computer labs, where the majority of instruction is provided through interaction with computer software. “The software provides examples, explanations, videos, opportunities to practice, and feedback on incorrect solutions” (Erin E. Krupa, 2014).

Also, college is expensive and nontraditional and underrepresented students have typically attended schools with low-achieving or failing students (Wimberly & Noeth, 2005). Despite financial incentives provided by federal and state governments for these students, they are underrepresented at colleges with higher-achieving students (Adelman, 2006; Bulger, S., & Watson, 2006). One study showed that the pass rate for these students increases 5% for a computer-based Intermediate Algebra course versus face-to-face (Erin E. Krupa, 2014).

Today, colleges are trying to use new methods of information technology to enhance the process and put a new face on learning to draw in more students. Rensselaer Polytechnic Institute provided $200,000 grants to thirty institutions to help these institutions redesign technological instruction for learning managed by Rensselaer . At the thirty institutions, replacing the traditional lecture format with technology reduced cost while also not lowering learning outcomes or achievement (Massy & Zemsky, 1995) According to Twigg, traditional math students do not actually do the problems; they do not spend enough time engaging with the material. She believes that information technology has matured enough to be the answer to learning (Carlo A. Twigg, 2011). Also, Krupa believes computer based learning is now the answer to progress over traditional learning for remedial courses (Erin E. Krupa, 2014) . Sixty percent (60%) of attrition occurs in the first two years of college (Mehaffy, 2012), thus if technology can help students succeed in their initial years (Complete College America, 2012, 2016), then this may also reduce the dropout rate.

The University of New Mexico main campus adopted the emporium idea in the Fall 2012, and created the Math Mαll Lab which replaced face-to-face instruction with computer-based instruction for Intermediate Algebra. Note that most sections of College Algebra, MATH 121, are taught face-to-face, but can optionally be taken as a computer-based course (The University of New Mexico, 2017). Furthermore, University College took over the Introductory Studies – Mathematics (ISM 100) course in Fall 2014 calling it Quantitative Reasoning (UNIV 102), and revising it again in Fall 2015 to Foundational Math (UNIV 103) using the computer-based (ALEKS) instruction. Now the UNIV 103 design is very similar to the sequence of MATH 101-102-103 in the Math Mαll.

1.7 Research Question

Given that we still aim to help every student succeed, even in the light of the massively positive effect ALEKS has had on the course's success rate, this thesis answers this question: Which factors influence student success in Intermediate Algebra, MATH 101-102-103? The goal is to develop a statistical model to help predict which students are most likely to benefit from early additional academic support in the class to improve their success, and therefore further improve the overall success rate in the course.

2 Methods

Our goal is to model student success in MATH 101 (defined either as their final letter grade, as a numeric GPA equivalent grade, or as a binary pass/fail) as a function of predictor variables for each student. We first introduce the variety of data sources then briefly describe each analysis method used.

2.1 Data sources

The data sources and variables available from each source are listed in Table 1. Data from the UNM Registrar (obtained Fall 2016 and updated in Fall 2017, http://oia.unm.edu/data-requests/data-request.html), public sources, and the ALEKS computer system were all joined together into a single large table for analysis. From the UNM Registrar, for all students who have taken MATH 101-102-103, we have student home addresses, high schools, test scores, math course grades, and details from their MATH 101-102-103 courses. From public sources obtained 2014 (The University of New Mexico, 2017; United States Census Bureau, 2010a, 2010b) we have county names for each NM postal code, as well as poverty data for each county. From ALEKS (obtained June 2017, via Harold D. Baker and Werner Garciano at ALEKS.com with IRB-exempt request) we have student responses for assessments, and every student response and request for explanation for practice questions, and have derived response time on practice questions.

By joining all of these data sources we have many questions we can answer about factors influencing student success. The postal (zip) codes from the home addresses are used to match with NM county poverty data, since most of our students are from NM. The zip code and NM poverty data help us understand the socioeconomic situation that the students are coming from (at the county level). The high school data was intended to understand whether students matriculated from some schools that struggled more than others, however we are already only looking at some of the lesser-prepared students from all schools. The test scores data provided ACT composite and component scores for each student, these scores are used partly for student math placement. The final grades for students in all math courses help us understand the historical trends for the MATH 120 and MATH 101-102-103 courses. The MATH 101-102-103 data provide the final grades and student characteristics for all students who have ever taken this course sequence.

The ALEKS data is very rich and was challenging to reshape into a usable format and to develop meaningful summaries (or features) for analysis. The data only cover the 786 students who have been enrolled under the new student interface (NSI) from 2016-07-05 to 2017-06-01 (primarily from Fall 2016). Some of these student records could not be easily matched with registrar data. While students are asked to enter their UNM ID number, many were entered incorrectly; furthermore, many students didn’t sign up with their UNM email address. We went through any students who didn’t match up and resolved all that we could from the UNM directory information.

The ALEKS data includes two types: practice questions and assessments. For practice questions we know the exact time the question was answered, whether they asked for help, how they answered, and whether they tried the question again immediately or at a later date. With some exceptions, the time between two questions can be used as the length of time spent on the second of these questions. Assessments are exams that include a set of questions, from each we have which questions were answered correctly, incorrectly, or "I don't know". From assessments we can summarize the proportion of questions correct, as well as many other features (which we didn't) such as the number of questions from each topic, etc.

All of these data are joined by the student ID information and analyzed rather pragmatically to help answer our primary research question: Which factors influence student success in Intermediate Algebra, MATH 101-102-103?

Table 1 Data dictionary of data sources and variables

Source	Filename	Variable	Type	Values	Comments
UNM_Registrar	Home Address.xlsx	ID	num	123456789	9-digit UNM ID
UNM_Registrar	Home Address.xlsx	NAME	char	Last, First M.	
UNM_Registrar	Home Address.xlsx	STREET_LINE1	char		
UNM_Registrar	Home Address.xlsx	STREET_LINE2	char		often . if nothing
UNM_Registrar	Home Address.xlsx	CITY	char		
UNM_Registrar	Home Address.xlsx	STATE_PROVINCE_DESC	char		
UNM_Registrar	Home Address.xlsx	POSTAL_CODE	char	12345-1234	
UNM_Registrar	Home Address.xlsx	NATION_DESC	char		. if USA, otherwise other country
UNM_Registrar	High Schools.csv	ACADEMIC_PERIOD	num	yyyys0	year, semester (1=Spring, 8=Fall, 6=Summer), 0 filler?
UNM_Registrar	High Schools.csv	ID	num	123456789	9-digit UNM ID
UNM_Registrar	High Schools.csv	NAME	char		
UNM_Registrar	High Schools.csv	INSTITUTION	num		High School Number
UNM_Registrar	High Schools.csv	INSTITUTION_DESC	char		High School Name
UNM_Registrar	Test Scores.csv	ACADEMIC_PERIOD	num	yyyys0	year, semester (1=Spring, 8=Fall, 6=Summer), 0 filler?
UNM_Registrar	Test Scores.csv	ID	num	123456789	9-digit UNM ID
UNM_Registrar	Test Scores.csv	NAME	char	Last, First M.	
UNM_Registrar	Test Scores.csv	TEST	char		Test label
UNM_Registrar	Test Scores.csv	TEST_DESC	char		Test name
UNM_Registrar	Test Scores.csv	TEST_SCORE	num		
UNM_Registrar	Test Scores.csv	TEST_DATE	date	MM/DD/YYYY	
UNM_Registrar	All Math Courses.xlsx	ID	num	123456789	9-digit UNM ID
UNM_Registrar	All Math Courses.xlsx	ACADEMIC_PERIOD	num	yyyys0	year, semester (1=Spring, 8=Fall, 6=Summer), 0 filler?
UNM_Registrar	All Math Courses.xlsx	ACADEMIC_PERIOD_DESC	char	Semester YYYY	
UNM_Registrar	All Math Courses.xlsx	COURSE_IDENTIFICATION	char	MATHnnn	
UNM_Registrar	All Math Courses.xlsx	COURSE_SECTION_NUMBER	num		
UNM_Registrar	All Math Courses.xlsx	COURSE_REFERENCE_NUMBER	num	12345	UNM 5-digit number
UNM_Registrar	All Math Courses.xlsx	COURSE_CREDITS	num	1, 2, 3	
UNM_Registrar	All Math Courses.xlsx	FINAL_GRADE	char	A+-, ..., F, W, NC, CR	
UNM_Registrar	All Math Courses.xlsx	INSTRUCTION_DELIVERY_MODE	char	., ONL, ENH, CORR, ...	
UNM_Registrar	All Math Courses.xlsx	INSTRUCTION_DELIVERY_MODE_DESC	char	., Web Enhanced, ...	description of instruction delivery
UNM_Registrar	All Math Courses.xlsx	PRIMARY_INSTRUCTOR_ID	num	123456789	Instructor 9-digit UNM ID
UNM_Registrar	All Math Courses.xlsx	PRIMARY_INSTRUCTOR_FIRST_NAME	char	First Name	Instructor First Name

Source	Filename	Variable	Type	Values	Comments
UNM_Registrar	All Math Courses.xlsx	PRIMARY_INSTRUCTOR_LAST_NAME	char	Last Name	Instructor Last Name
UNM_Registrar	Math101-102-103_Students.csv	ACADEMIC_PERIOD	num	yyyys0	year, semester (1=Spring, 8=Fall, 6=Summer), 0 filler?
UNM_Registrar	Math101-102-103_Students.csv	COURSE_IDENTIFICATION	char	MATHnnn	
UNM_Registrar	Math101-102-103_Students.csv	COURSE_SECTION_NUMBER	num	1, 2, 3, ...	
UNM_Registrar	Math101-102-103_Students.csv	COURSE_REFERENCE_NUMBER	num	12345	UNM 5-digit number
UNM_Registrar	Math101-102-103_Students.csv	INSTRUCTION_DELIVERY_MODE_DESC	char	., Web Enhanced, ...	description of instruction delivery
UNM_Registrar	Math101-102-103_Students.csv	PRIMARY_INSTRUCTOR_ID	num	123456789	Instructor 9-digit UNM ID
UNM_Registrar	Math101-102-103_Students.csv	PRIMARY_INSTRUCTOR_FIRST_NAME	char	First Name	Instructor First Name
UNM_Registrar	Math101-102-103_Students.csv	PRIMARY_INSTRUCTOR_LAST_NAME	char	Last Name	Instructor Last Name
UNM_Registrar	Math101-102-103_Students.csv	ID	num	123456789	9-digit UNM ID
UNM_Registrar	Math101-102-103_Students.csv	NAME	char	Last, First M.	
UNM_Registrar	Math101-102-103_Students.csv	FINAL_GRADE	char	A+-, ..., F, W, NC, CR	
UNM_Registrar	Math101-102-103_Students.csv	STUDENT_LEVEL	char	UG, G	
UNM_Registrar	Math101-102-103_Students.csv	STUDENT_LEVEL_DESC	char	Undergraduate, ...	
UNM_Registrar	Math101-102-103_Students.csv	MAJOR	char	CRIM, ENGL, BUS, BIOL, ...	
UNM_Registrar	Math101-102-103_Students.csv	MAJOR_DESC	char	Biology, ...	
UNM_Registrar	Math101-102-103_Students.csv	COLLEGE	char	AS, UC, ...	
UNM_Registrar	Math101-102-103_Students.csv	COLLEGE_DESC	char	College of Arts and Sciences, ...	
UNM_Registrar	Math101-102-103_Students.csv	DEPARTMENT	char	123A	Department code
UNM_Registrar	Math101-102-103_Students.csv	DEPARTMENT_DESC	char	Sociology, ...	Department name
UNM_Registrar	Math101-102-103_Students.csv	STUDENT_CLASSIFICATION	char	UB, U2, U3, U4, ...	
UNM_Registrar	Math101-102-103_Students.csv	STUDENT_CLASSIFICATION_DESC	char	Freshmen, 1st Yr, 2nd Sem; â€¦	
UNM_Registrar	Math101-102-103_Students.csv	BIRTH_DATE	date	MM/DD/YYYY	
UNM_Registrar	Math101-102-103_Students.csv	GENDER_DESC	char	Male, Female	
UNM_Registrar	Math101-102-103_Students.csv	IPEDS_VALUES_DESC	char	White, Hispanic,	

Source	Filename	Variable	Type	Values	Comments
				American Indian, ...	
Public	zipcode_county.csv	POSTAL_CODE	num		
Public	zipcode_county.csv	City	char		
Public	zipcode_county.csv	County	char		
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	area	num		County number
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	County	char		County name
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	periodyear	num	2014	All 2014
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	Homeowner vacancy rate	num	percent	Vacancy Status
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	Rental vacancy rate	num	percent	Vacancy Status
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	Owner-occupied	num	percent	Occupancy Status
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	Renter-occupied	num	percent	Occupancy Status
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	Utility gas	num	percent	Main source of heating
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	Electricity	num	percent	Main source of heating
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	Wood	num	percent	Main source of heating
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	Lacking complete plumbing facilities	num	percent	Poverty
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	No telephone service available	num	percent	Poverty
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	rentover35	num	percent	Poverty
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	est_percent	num	percent	Poverty
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	child_percent	num	percent	Poverty
bber.unm.edu/ county-profiles	NMCensusPoverty HousingCharacteristics_DP04.csv	fam_percent	num	percent	Poverty
ALEKS	aleks_student.csv	database_code	char	home_highed_db_xkc9vem	All the same
ALEKS	aleks_student.csv	student_ID	char		

Source	Filename	Variable	Type	Values	Comments
ALEKS	aleks_student.csv	email	char	netid@unm.edu	
ALEKS	aleks_student.csv	ID	num	123456789	9-digit UNM ID
ALEKS	aleks_student.csv	class_code	char		
ALEKS	aleks_student.csv	course_name	char	begint.1	All the same
ALEKS	aleks_practice.csv	student_ID	char		
ALEKS	aleks_practice.csv	datetime	date		
ALEKS	aleks_practice.csv	event_type	char	C, W, E, S, F	either correct answer , wrong/IDontKnow answer , explanation lookup , final correct answer (for succeed), or final incorrect answer (for fail)
ALEKS	aleks_practice.csv	item_ID	char	+geom807	+ means trying again on their own
ALEKS	aleks_practice.csv	plus	char	+ or NA	
ALEKS	aleks_practice.csv	subject	char	geom	
ALEKS	aleks_practice.csv	Qnum	num	807	
ALEKS	aleks_assessment.csv	student_ID	char		
ALEKS	aleks_assessment.csv	reason	char	init, progress, swclass, reqany, goal, time	
ALEKS	aleks_assessment.csv	type	char	init, progress, graded	
ALEKS	aleks_assessment.csv	date	date		
ALEKS	aleks_assessment.csv	ans_cor	char	geom808,alge482,alge471,...	list of questions answered correctly
ALEKS	aleks_assessment.csv	ans_inc	char	geom808,alge482,alge471,...	list of questions answered incorrectly
ALEKS	aleks_assessment.csv	ans_idk	char	geom808,alge482,alge471,...	list of questions indicated I don't know
ALEKS	aleks_assessment.csv	ind_ans_cor	char	geom808,alge482,alge471,...	ignored ALEKS trial questions
ALEKS	aleks_assessment.csv	ind_ans_inc	char	geom808,alge482,alge471,...	ignored ALEKS trial questions
ALEKS	aleks_assessment.csv	ind_ans_idk	char	geom808,alge482,alge471,...	ignored ALEKS trial questions
ALEKS Derived variables	Several				

2.2 Data decisions

While an ACT score between 18 and 21 is required to enter MATH 101-102-103, a very small proportion of students with a higher ACT score have taken the class in order to earn one easy credit. As of Fall 2017, students are no longer allowed to take these courses for credit when they have higher ACT scores (The University of New Mexico, 2017). Thus, only students with the appropriate range of ACT scores have been included (Figure 8). Letter grades were converted to their GPA equivalent through the following table (Table 2).

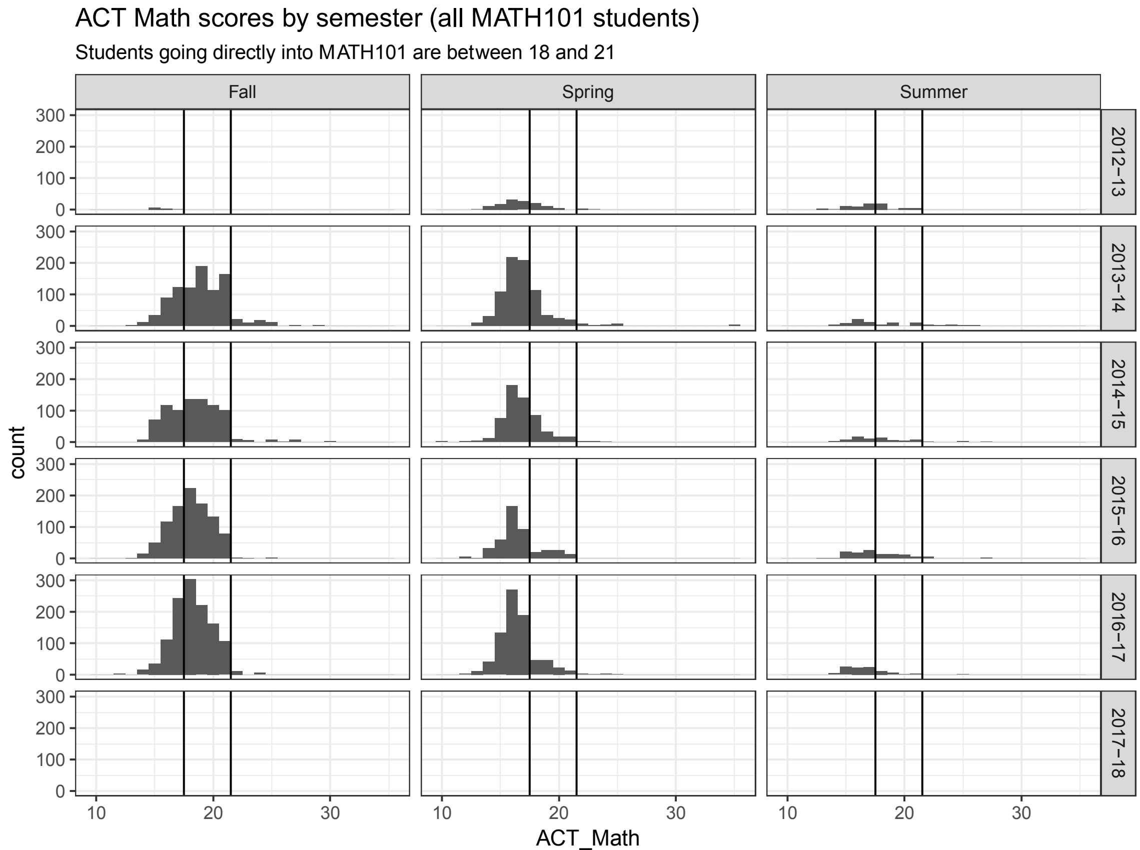

Figure 8 Historic of range ACT Math Score with cut off from 18-21. Note that the Spring MATH 101 students have lower ACT Math scores because they will have taken UNIV 103 in the previous Fall.

Table 2 Convert final letter grade to GPA

FINAL_GRADE	A+	A	A-	B+	B, CR, PR	B-	C+	C
FINAL_GPA	4.33	4	3.67	3.33	3	2.67	2.33	2

FINAL_GRADE	C-	D+	D	D-	F, WF, IF, NC, I, INC	AUD, WP, WNC, W, XW, RS, NR, .
FINAL_GPA	1.67	1.33	1	0.67	0	NA

[redacted]ALEKS Assessment and Practice

We processed both the practice and assessment ALEKS datasets to produce useful and interpretable features for analysis.

2.2.1.1 ALEKS Assessment summaries

The periodic assessments evaluate a student's knowledge. Assessments are given on the first day of class (baseline) and thereafter after 5 hours of practice or 20 practice topics, whichever comes first. We will focus on the initial (baseline) assessment because we're interested in early-semester intervention. For the assessment, we summarize the number questions answered correctly, incorrectly, marked "I don't know", then we also total the number of questions and calculate the proportion correct. ALEKS includes a few "test" quality control questions in each assessment which are not counted for credit and these were ignored.

2.2.1.2 ALEKS Practice question summaries

Students spend most of their time learning how to solve types of problems and practicing them. Practice questions leading up to an assessment are either marked "correct" or "succeed"; "wrong" or "fail"; or "explanation" when a student asks for further help. A record is recorded when the student clicks to grade their response to the question or clicks for explanation, the question item ID (subject and question number) and the date/time are included. Each response indicates a decision after engaging with the content in ALEKS. A typical sequence of data appears in Table 3. Using the algorithm below for determining question timing, we derived summary metrics for the mean time to answer a question correctly or incorrectly, reading explanations, the overall average time between clicks, and the ratio of the average time for a correct vs incorrect response.

Table 3 ALEKS practice question timing algorithm

datetime	event _type	item_ID	Time	add_time	Comment
…					Note: When you get a question correct or ask for explanation, the next question will differ in the variable values.
12/13/2016 5:10	S	alge761	…		
12/13/2016 5:14	C	alge180	4		From previous answer
12/13/2016 5:15	C	alge180	1		
12/13/2016 5:16	E	alge180	1		Half of time from previous answer to next answer
12/13/2016 5:17	C	alge180	1		The second half of the time from the previous answer to this answer
12/13/2016 5:18	S	alge180	1		
12/13/2016 6:03	W	alge739	NA		Large gaps in time are replaced with the average length of time for completing a question
12/13/2016 6:04	W	alge739	average wrong + 1	1	Consecutive wrong have same variable values. Use only the last "wrong" for timing in consecutive wrong answers.
12/13/2016 6:04	E	alge739	-		Combine consecutive explanations
12/13/2016 6:05	E	alge739	1.5		Half of time from previous answer to next answer
12/13/2016 6:07	W	alge739	-		Use only the last "wrong" for timing in consecutive wrong answers.
12/13/2016 6:07	W	alge739	1.5		The second half of the time from the previous answer to this answer
12/13/2016 6:07	E	alge739	2		
12/13/2016 6:11	C	alge739	2		
12/13/2016 6:14	C	alge739	3		
12/13/2016 6:19	S	alge739	5		
12/13/2016 6:26	W	alge762	7		
12/13/2016 6:26	C	alge762	7		Same values in this question, thus, it took a total of 7 minutes to get it wrong and then correct it

datetime	event_type	item_ID	Time	add_time	Comment
12/13/2016 6:33	E	alge762	6		Half of time from previous answer to next answer
12/13/2016 6:38	W	alge762	6		The second half of the time from the previous answer to this answer
12/13/2016 6:39	C	alge762	6+1 = 7		Same values in this question
12/13/2016 6:43	C	alge762	4		
12/13/2016 6:47	S	alge762	4		
…					

A student spends time working on each question and either provides a correct or incorrect answer, or asks for an explanation. After reviewing an explanation, they return to the same question but with different numeric values and can again either answer the question or ask for explanation again. If a question is answered incorrectly, they have another chance to correct their work and answer the question with the same numeric values. If a question is answered correctly and they are given the same item ID, then it's the same type of question but with different numeric values. To determine the length of time students spend on asking for explanation versus the time it takes to answer a question incorrectly or correctly, we developed the following rules.

Rename "S" to "C", and "F" to "W" since those are the same except that "S" and "F" indicate the last attempt for a particular item.

For each student …

1. If this is the first record, AND response is "C" or "W", then insert the time as -9 and return later.
2. If the previous record is more than 10 minutes before, insert the time as -9 and return later. The gap is probably between work sessions. We will

impute this time as the average time for this response type (“C”, “W”, or “E”).

For each item, …

3. Collapse consecutive explanations (E) to first one (if same item ID) because students may ask for explanation (E) multiple times in a row without answering the question.
4. Collapse consecutive wrong answers (W) to last one (if same item ID) because students may answer the same question incorrectly multiple times.
    a. If there was a 10-minute gap for an NA, then we want to add the time difference between the first and last consecutive “W” to the average value that we’ll impute at the end. So, create an “add_time” column with that consecutive “W” time difference, put this value on the last row. Do this before collapsing rows.
5. When an answer is provided, look back or forward to answers to determine a starting point for timing.
    a. If this response is an “E”, look forward.
        i. If the next is a different item ID and “E”
            1. ignore it.
        ii. If the next is an “C”, then calculate the time between the previous result (“C” or “W”) and the next “C” and allocate 1/2 to each.

iii. If the next is an “W”, then calculate the time between the previous result (“C” or “W”) and the next “W” and allocate 1/2 to each.

b. If this response is a “W”, look back.

    i. If the previous is a “C”, calculate the time between the current result and the previous.

    ii. (If the previous is a “E”, already done.)

c. If this response is a “C”, look back.

    i. If previous is a “C”, then calculate the difference in time (this minus previous).

    ii. (If previous is a “E”, already done.)

    iii. If previous is a “W”, then calculate the difference in time (this minus previous) plus the time for the previous “W”. This indicates they have been continuing to work on this problem and eventually got it right after at least one wrong answer.

6. Impute missing NA values with the average length of time this student took on this item response type.

    a. Summarize the mean times for “E”, “C”, and “W”

    b. Fill in the NAs with these mean times for each type.

7. Make “add_time” additions to those 10-minute gap consecutive “W” questions.

8. DONE.

2.3 Statistical methods

All analyses were carried out using R 3.4.2 with extensive use of the base, ggplot2, and randomForestSRC packages (Ishwaran, H., & Malley, 2014; Wickham, 2009).

[redacted] Contingency tables

The chi-squared test for homogeneity of population proportions will be used for high-level comparisons of observed proportions of success between groups (Erhardt, Bedrick, & Schrader, 2013). Assume that the data are independent samples from r populations (strata or groups), and that each individual is placed into one of c levels of a categorical variable ($c = 2$ for pass or fail). The data can be summarized as a $c \times r$ contingency table of counts, where the rows correspond to the samples, and the columns are the levels of the categorical variable. Under the null hypothesis, row proportions are equal between groups. The Pearson chi-squared test assesses the evidence against the null. Pearson residuals are helpful for indicating which table cells most contradict the null hypothesis.

[redacted] Multiple linear regression

Note that for our GPA data, the model assumptions for linear regression of normal residuals are rarely met, so the random forest algorithm will be preferred for understanding variable importance for prediction.

In linear regression, we are interested in developing a linear equation that best summarizes the relationship in a sample between the response variable Y and the predictor variable (or independent variable) X (Erhardt et al., 2013). The equation is also used to predict Y from X.

First, we give a brief review of simple linear regression (SLR). Given a data set of paired observations from n subjects, (X_i, Y_i), $i = 1, \ldots, n$, the SLR model is usually written as $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, where β_0 is called the intercept (the predicted value of Y when $X = 0$), β_1 is called the slope, and both of them are unknown constants, and ε_i is a random error component (the part of Y not explained by the model). The errors are assumed to be normal random variables with mean 0 and constant variance σ^2. Since SLR involves only one predictor variable, it is called a “simple” linear regression.

In general, the response variable Y may be related to k predictors, $X_1, X_2, \ldots, X_k$, so that $Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \cdots + \beta_k X_{k,i} + \varepsilon_i$. This is called a multiple linear regression model as more than one predictor variable is involved. The word "linear" is indicating that the model is linear in the parameters $\beta_0, \beta_1, \ldots, \beta_k$. The important objective of regression is estimating the unknown parameters in the regression model. In most situations, the response variable Y will be the student's final numeric GPA. The model assumes: 1. The average Y-value at a given X-value is linearly related to X. 2. The variation in responses Y at a given X value is constant. 3. The population of responses Y at a given X is normally distributed. 4. The observed data are a random sample.

[redacted] Logistic regression

Logistic regression is a generalized linear modeling approach that fits a regression model using predictors with a categorical response variable (Success vs Failure); this method, while it is both familiar and interpretable, is susceptible to overfitting (Christensen, 2006). Logistic regression was performed in R software using the standard package "stats" function "glm" with option "family=binomial(link=logit)".

[redacted] Random forest (RF)

To describe a random forest, we first must define a classification tree. A classification (decision) tree is a schematic, tree-shaped diagram used to determine a predicted outcome (the leaves) based on a series of branching decisions (branches). The tree is structured to show a hierarchy of choices leading to a classification. The example tree below (Figure 9) illustrates using ACT component scores to predict whether a student's final grade is an A versus a B-F.

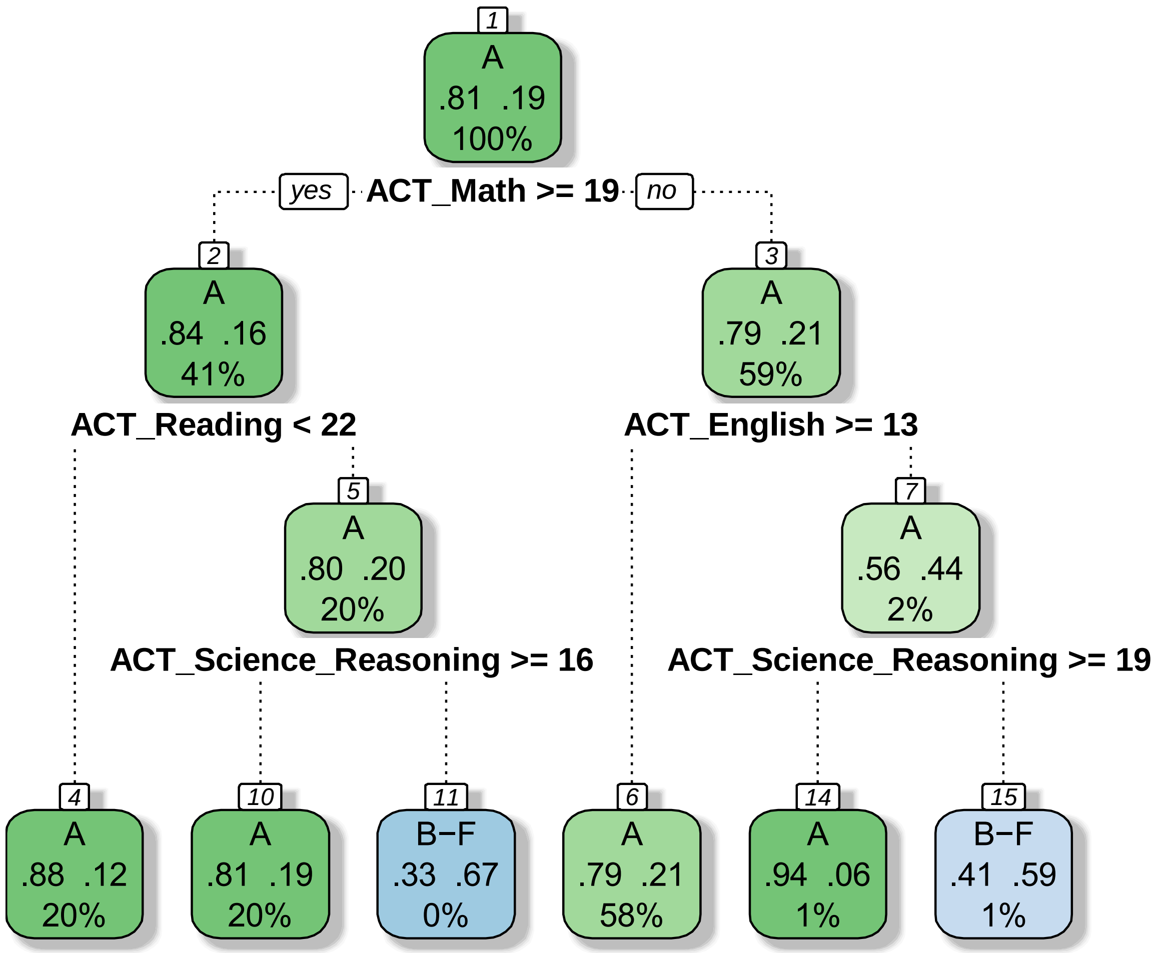

Figure 9 Predict final grade is A- or higher (A) versus B+ or lower (B-F) using ACT component scores.

Random forests (RF) are a supervised ensemble learning algorithm that is based on many classification trees (Breiman, 2001a, 2001b). Many classification trees (a "forest") are fit on bootstrapped samples of the original data. Each tree partitions the data based on a random subset of predictor variables in such a way as to try to get optimal separation of the response variable which is student's final GPA. RF provides a measure of variable importance for prediction accuracy, as well the marginal probability of group identity for values of each variable, and does not over fit (Breiman, 2001a, 2001b). Furthermore, RF can perform multiclass prediction, automatically employs external cross-validation by predicting the response variable based on trees estimated without that subject, and does not have distributional model assumptions, and is easy to implement. RF was performed in R software using package "randomForestSRC" function "rfsrc" with 1000 trees.

There are three important summaries that we'll use. For categorical response classification we'll have the confusion matrix (Table 4). For all response types we'll have two visual summaries. The first is the variable importance plot (VIMP) that indicates the relative importance of each variable in the accuracy of prediction or classification (Figure 10). The second are marginal effects plots that indicate the average effect of one predictor on the predicted response averaged over the values of the other predictors; these are ordered with respect to the VIMP (Figure 11).

Table 4 (Top) Random forest data and forest summaries. (Bottom) Confusion matrix for how the observed (true) labels were classified (predicted) along with the error rate.

Sample size:	3171,
Frequency of class labels:	2563, 608
Number of trees:	1000

	predicted		
observed	A	B-F	class.error
A	2502	61	0.0238
B-F	121	487	0.1990

Overall error rate: 5.74%

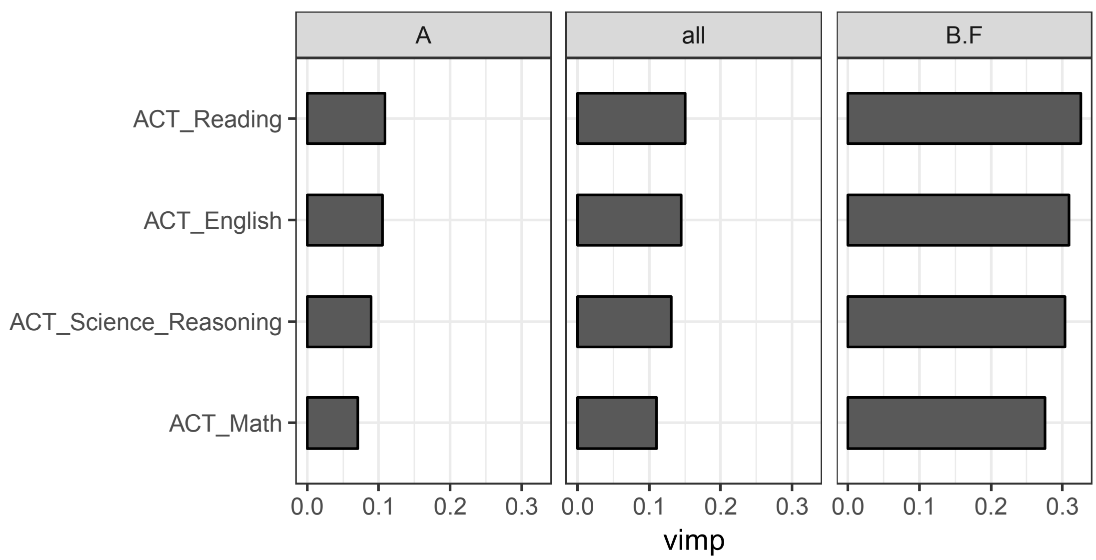

Figure 10 A variable importance plot. For both groups (all), Reading was the most important and Math the least important for classification accuracy.

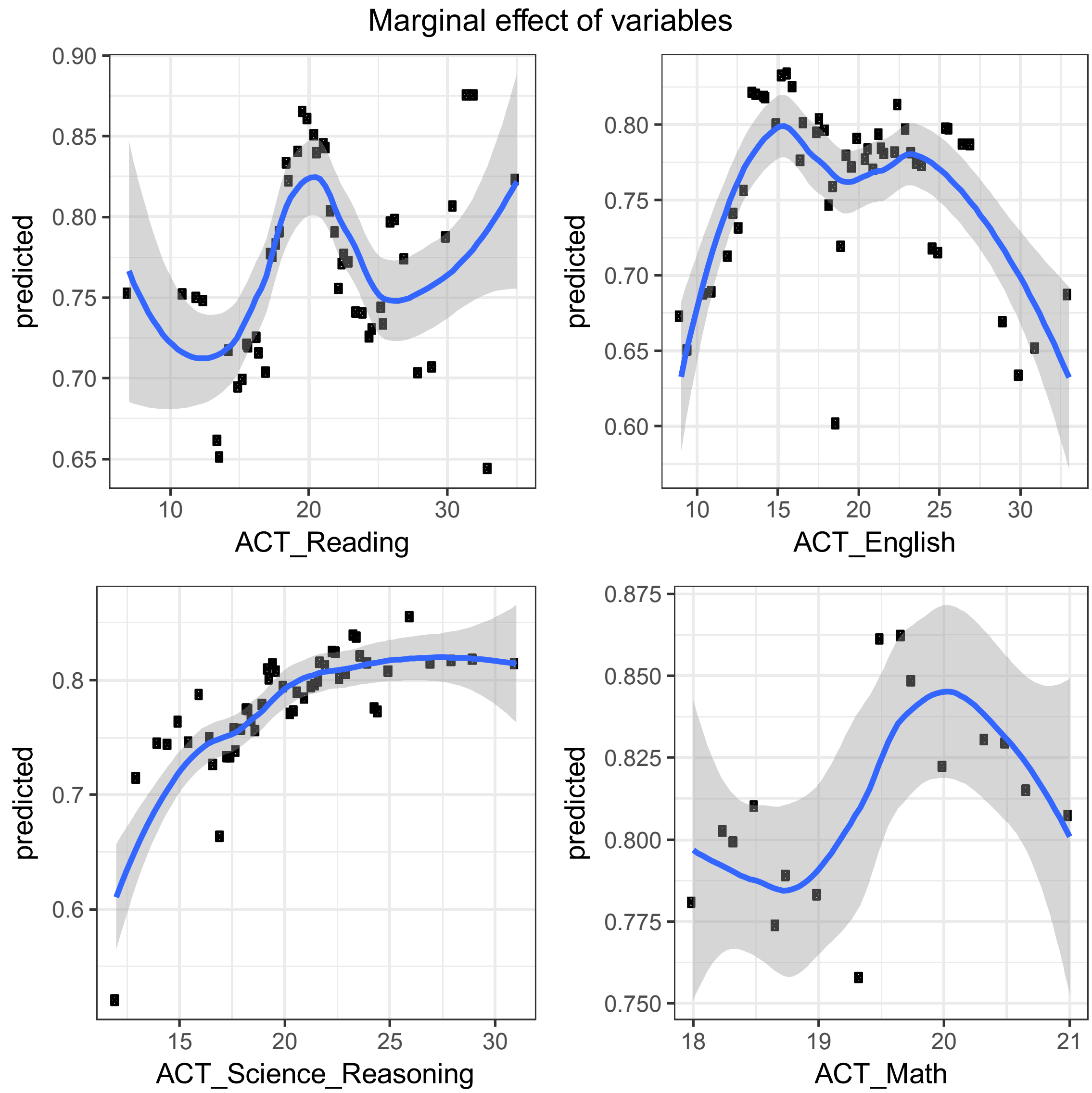

Figure 11 The marginal effects that indicate the average effect of one predicted response averaged over the values of the other predictor.

3 Results

3.1 UNM Registrar data

[redacted]Contingency tables

Success (passing with a C or better) is weakly related to student gender (Figure 13) and college they are enrolled in (Figure 12), and we have insufficient evidence that race is a factor (Figure 14). However, second-semester freshman taking MATH 101 do better (and probably most are coming from UNIV 103), while students waiting until their final years fail more often than their first- and second-year peers (Figure 15).

Figure 12 Success GPA of students in MATH 101 for different colleges

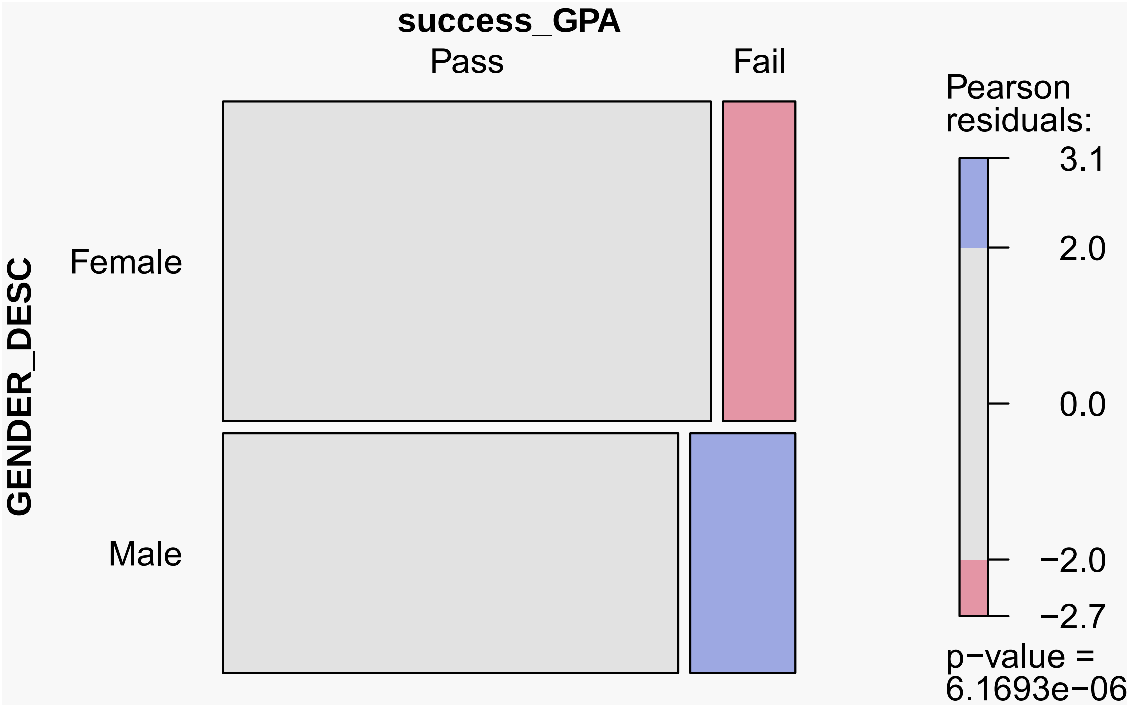

Figure 13 Success GPA of students in MATH 101 by gender

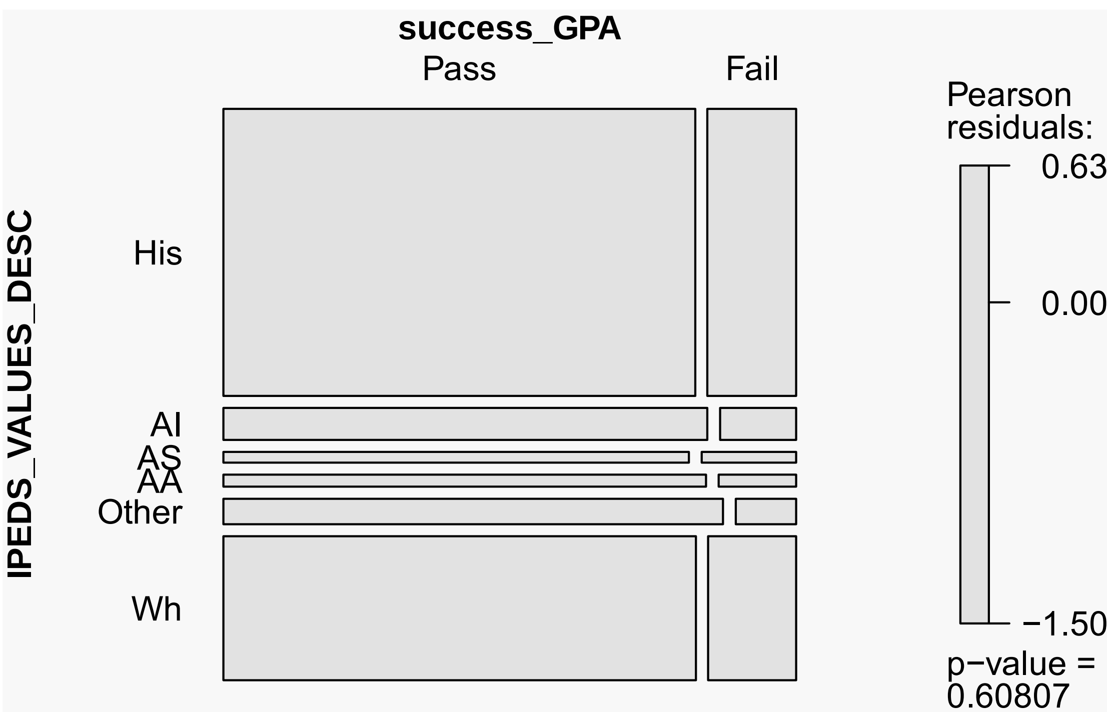

Figure 14 Success GPA of students in MATH 101 by race

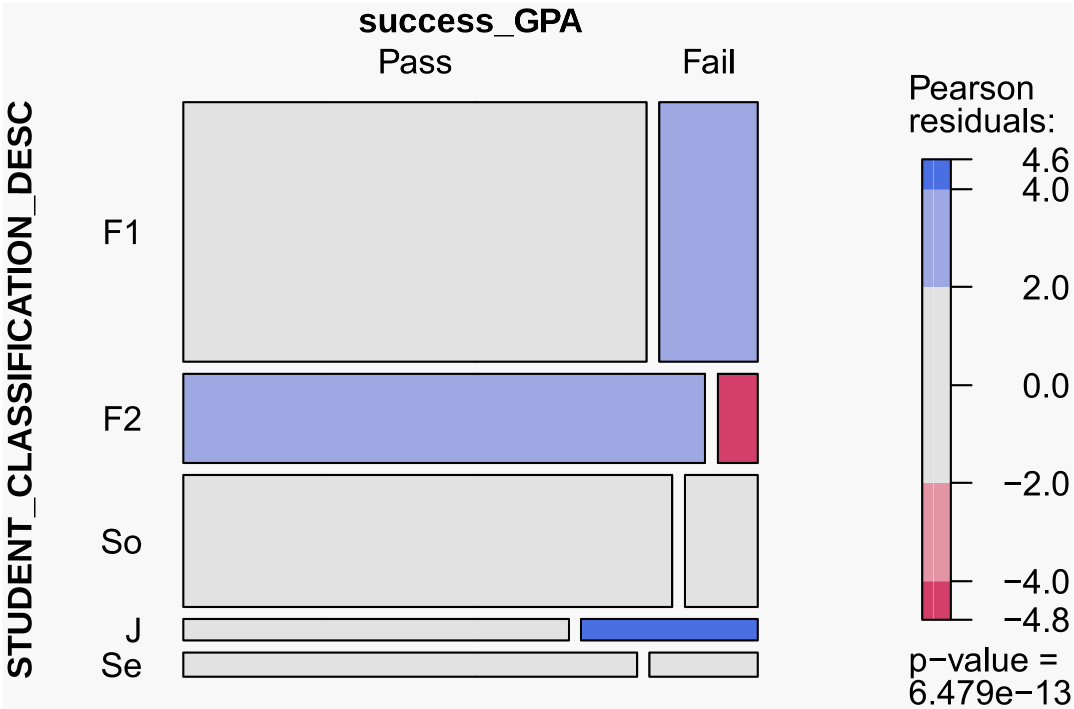

Figure 15 Success GPA of students in MATH 101 by year in college

[redacted]Multiple Linear Regression

Figure 16 illustrates that, marginally, no ACT or poverty-related variable has much of a relationship with the final GPA. Since the data will not meet the normality assumptions for multiple linear regressions, we will use a random forest to model the data.

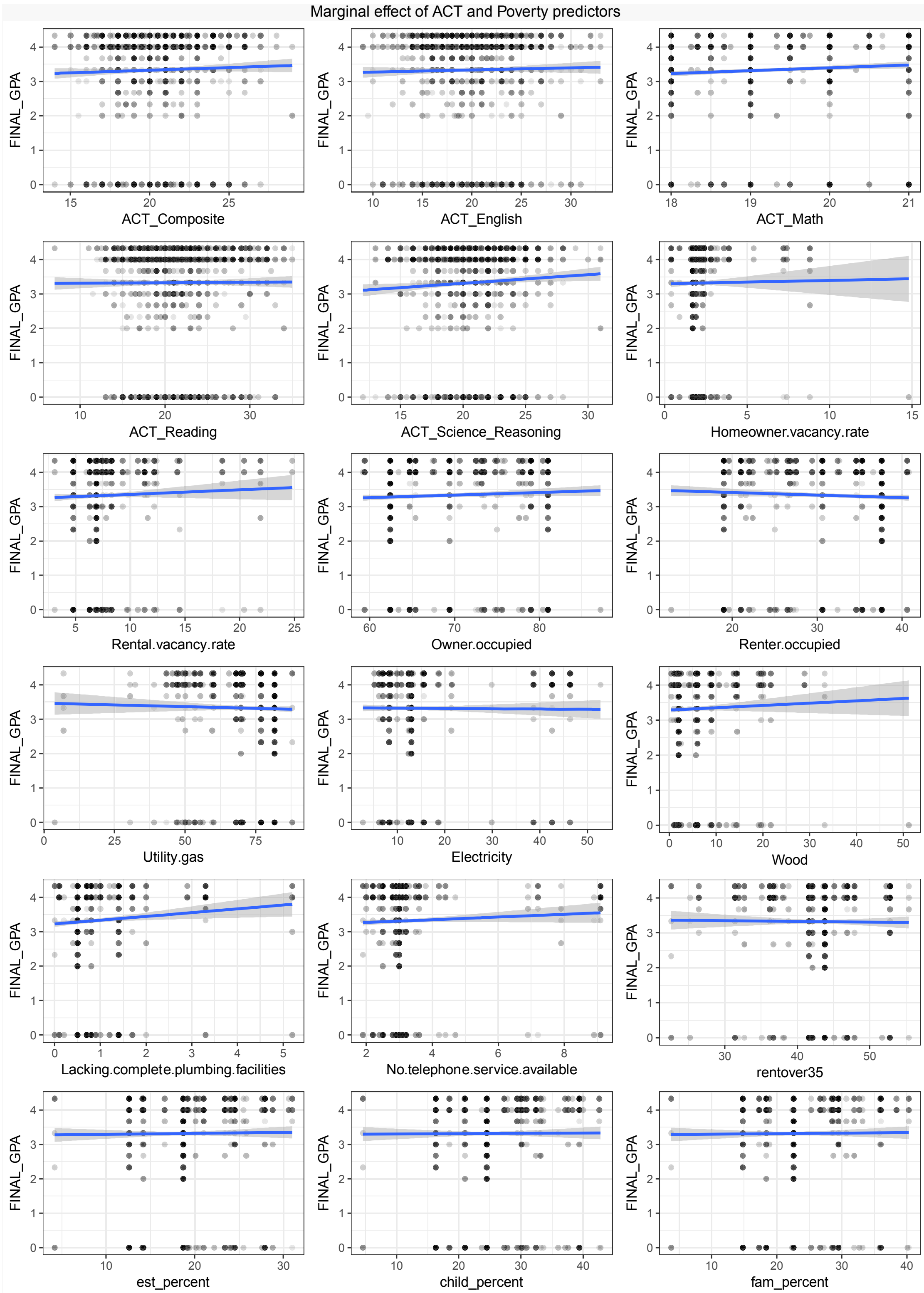

Figure 16 Final GPA versus ACT test scores and poverty variables

[redacted] Random forest

3.1.3.1 ACT scores predicting Final GPA

ACT scores explain roughly half (47.06%) of the variance in final GPA among MATH 101 students. The Reading, Composite, English, and Science Reasoning scores are more important than the Math score (Figure 17), though recall that we are only considering the subset of students with Math ACT scores in the range 18-21. The marginal patterns are not easily interpreted (Figure 18) largely because we are looking only at a subset of the students, this constrains the range of ACT values.

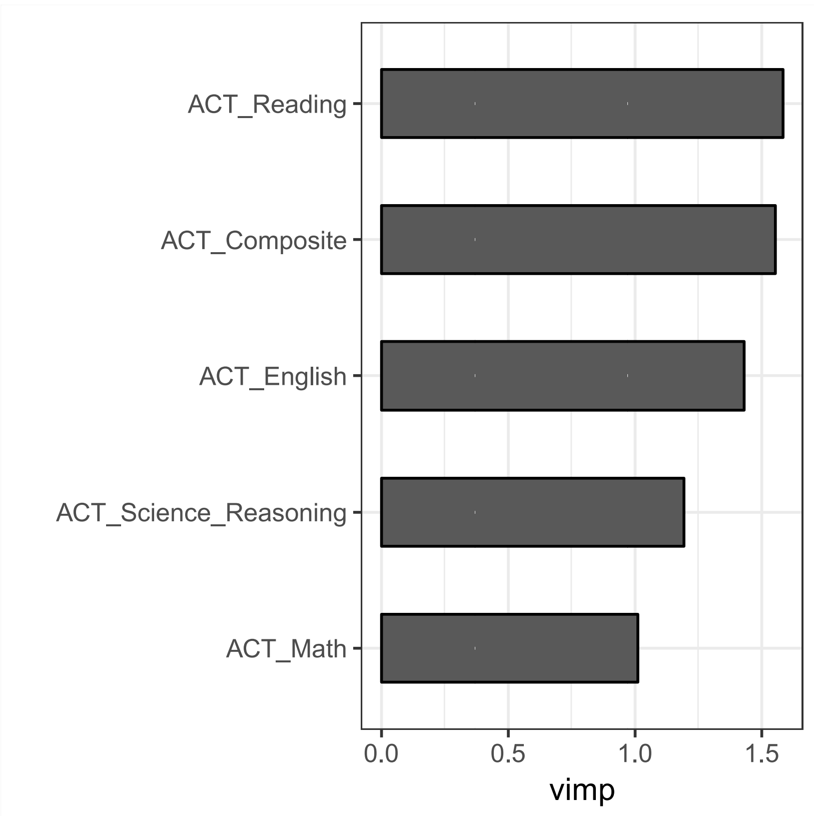

Figure 17 Variable importance for predicting final GPA from ACT scores

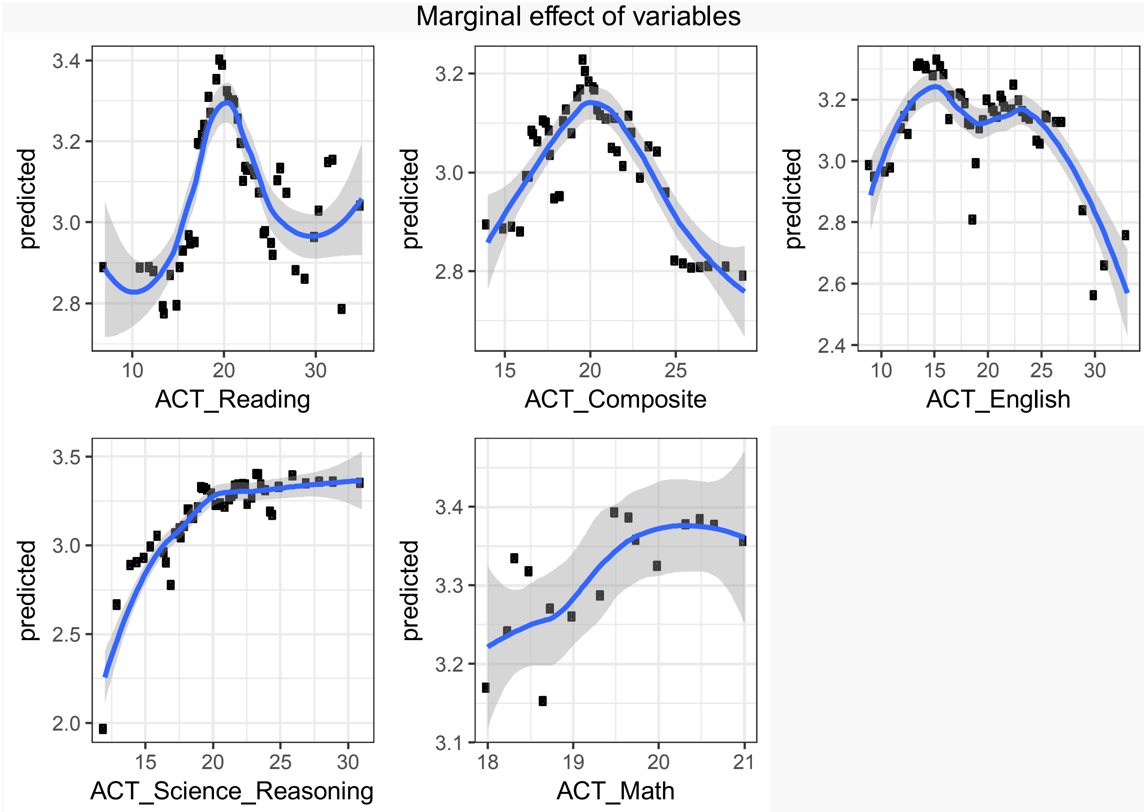

Figure 18 Marginal effects for predicting final GPA from ACT scores

3.1.3.2 NM Poverty data predicting Final GPA

The county-level poverty variables explain almost no variance (0.54%) in final GPA (results not shown). Thus, a student's GPA in MATH 101 appears not to depend on the poverty status of the county of New Mexico they come from.

3.1.3.3 ACT and NM Poverty data predicting Final GPA

Combining poverty variables to the ACT scores does not improve prediction over ACT alone (49.82% variance explained vs 47.06%). Figure 19 shows how much more important the ACT variables are over the poverty variables.

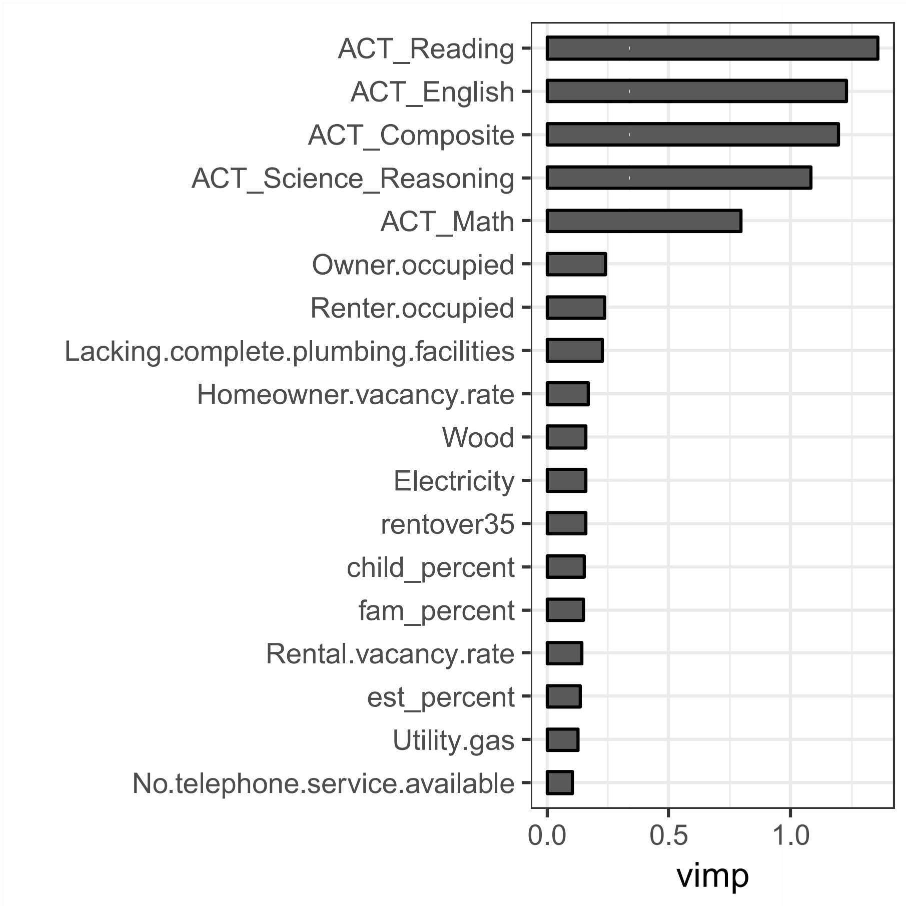

Figure 19 Variable importance for predicting final GPA from ACT scores and poverty variables

3.2 ALEKS data

The assessment and practice question summaries were analyzed in several ways.

[redacted]Logistic regression

We analyze the final grade in the course as pass/fail. The students who have success = "Fail" received an "NC", while "Pass" is any passing grade.

Note that in Table 5 we see an anomaly that students who failed overwhelmingly had initial assessments with 29 questions, while students who passed had fewer questions. In an attempt to reproduce this result, both Linh Ward and Dr. Srini Vasan took the initial assessment mimicking students who were "underprepared" and "over-prepared" for the course. Linh scored 10/26 and 25/30, while Srini scored 15/30 and 29/29. The total numbers of questions were 26 and 30 for the underprepared and 30 and 29 for the over-prepared. We could not reproduce the 29-question means "fail" result in the table, though it's possible their behavior didn't represent how students behave (for example, answering questions incorrectly without waiting or thinking about them). Therefore, in the analyses that follow, we exclude the total number of questions for the assessment since it's a serendipitous and strong classifier for Passing. All of this is to say that ALEKS may behave in mysterious ways, and that way may change from semester to semester.

Table 5 Number of students passing or failing MATH 101 conditional on the number of baseline assessment questions they were asked.

Number of questions	8	13	14	16	17	18	19	20	21	22	23	24	29
Pass	2	6	4	1	4	2	8	10	19	8	11	272	3
Fail	0	0	0	0	0	0	0	0	0	3	0	1	**44**

3.2.1.1 Assessment 0 predicting MATH 101 Success

We start with a model to predict student success in the course (passing with a C or better) using summaries from the initial assessment. Because

ALEKS adapts the assessment based on how the student answered the previous questions, it is difficult to interpret the “number of questions answered” variables (Figure 20 and Figure 21); they all predict that the students are less likely to succeed when more questions are asked/answered (as discussed in reference to Table 5). However, we do see an intuitive pattern with the proportion of questions answered correctly. Furthermore, the overall error rate is 1.26%, so the model predicts extremely well (Table 6).

Table 6 Random Forest summaries for predicting MATH 101 Success from Assessment 0

```
                         Sample size: 398
           Frequency of class labels: 350, 48
                     Number of trees: 1000

Confusion matrix:

          predicted
  observed Pass Fail class.error
      Pass  348    2      0.0057
      Fail    3   45      0.0625

    Overall error rate: 1.26%
```

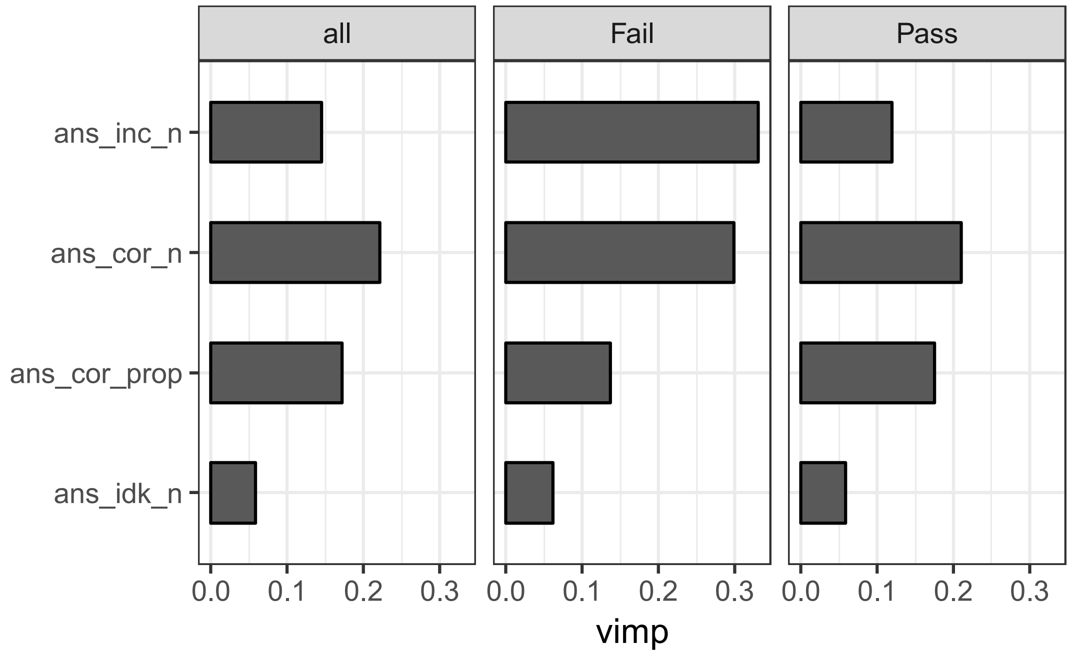

Figure 20 Random Forest VIMP for predicting MATH 101 Success from Assessment 0

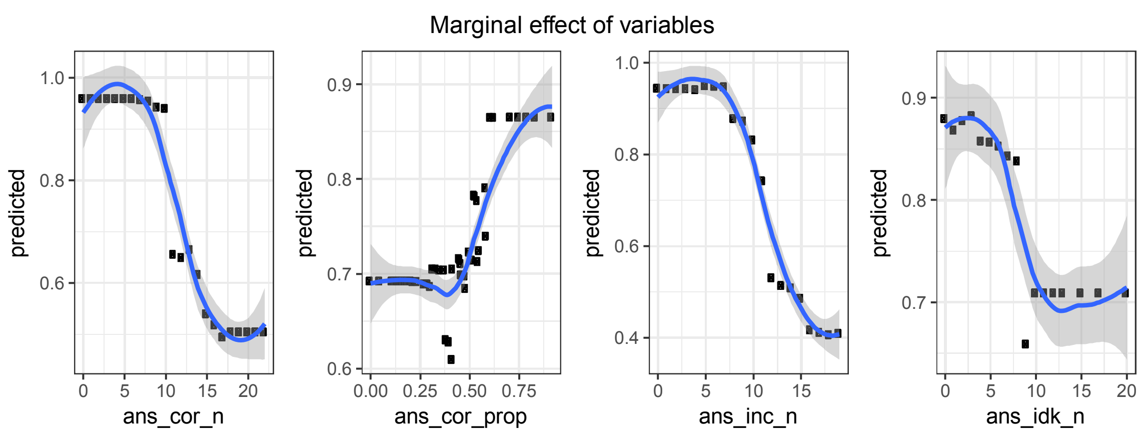

Figure 21 Random Forest Marginal Effect Plot for predicting MATH 101 Success from Assessment 0

3.2.1.2 Practice question times predicting MATH 101 Success

The mean times for different outcomes while practicing questions are also important features for predicting success (Figure 22). The marginal effects plot shows that spending about 2 minutes per question is associated with students who are most likely to succeed in the class (Prac_Qtime_mean_W, Prac_Qtime_mean_all, Prac_Qtime_mean_C); less than 90 seconds per question is typically associated with failing (Figure 23). Students who spend more time on reading explanations are also more likely to succeed (Prac_Qtime_mean_E). Students who succeed spend slightly less time answering a question correctly than incorrectly (Prac_Qtime_C_W_ratio). Furthermore, the overall error rate is 2.73%, but does not predict as well as the baseline assessment model (Table 7).

Table 7 Random Forest summaries for predicting MATH 101 Success from Practice question timing

```
                         Sample size: 366
           Frequency of class labels: 319, 47
                     Number of trees: 1000

Confusion matrix:

          predicted
  observed Pass Fail class.error
      Pass  317    2      0.0063
      Fail    8   39      0.1702

    Overall error rate: 2.73%
```

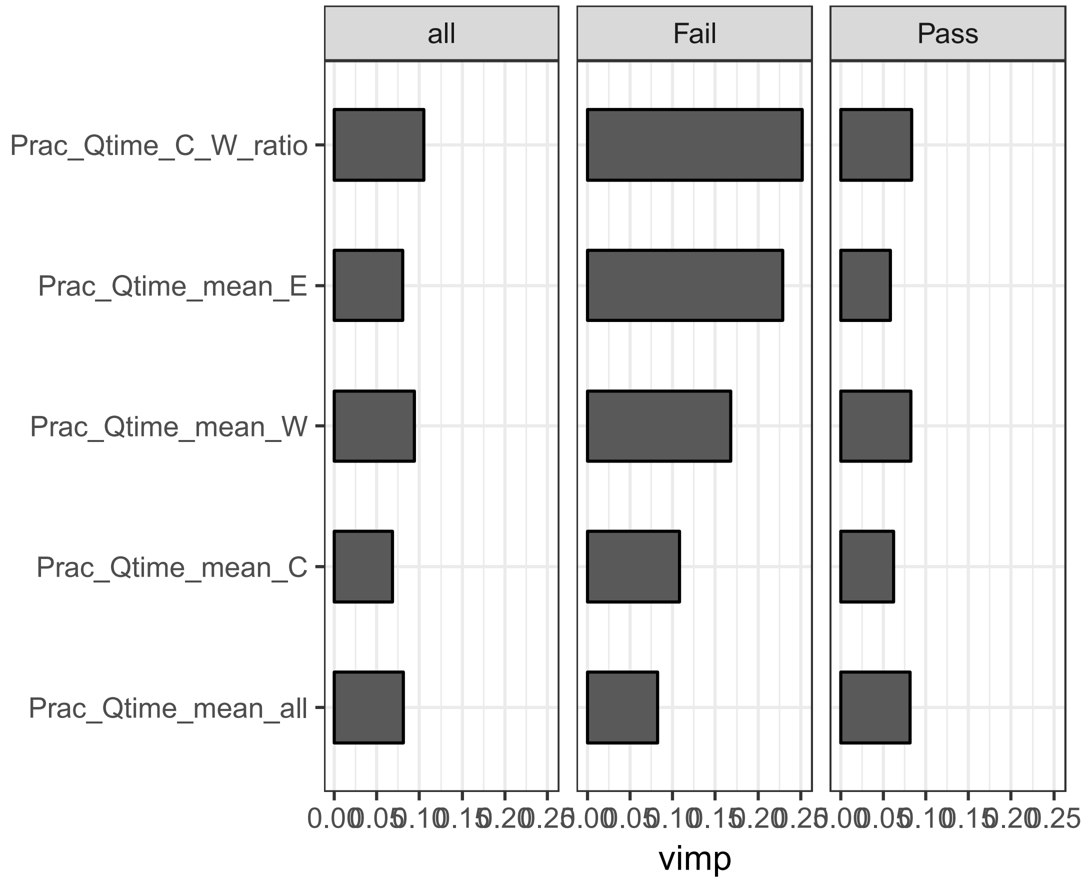

Figure 22 Random Forest VIMP for predicting MATH 101 Success from Practice question timing

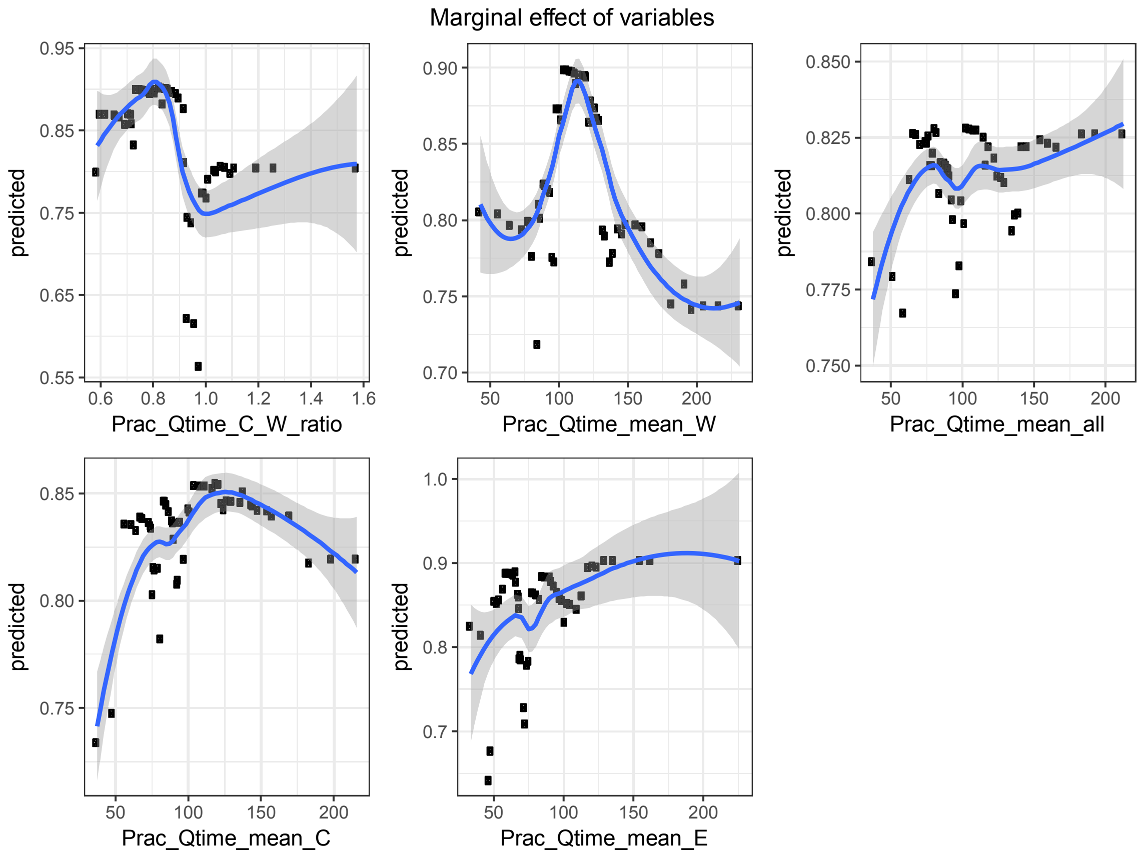

Figure 23 Random Forest marginal effects plot for predicting MATH 101 Success from Practice question timing

3.2.1.3 ACT Scores, demographics, and Assessment 0 times predicting MATH 101 Success (preferred model)

The baseline assessment variables are more important than the Registrar ACT and demographic variables. The ACT and demographic variables help lower the prediction error from 1.26% to 1.01% (Table 8). The marginal effects follow the same patterns for the assessment variables as they did without also considering the registrar variables. This is our best model with an overall error rate of 1.01% and stresses the marginal importance of basic literacy skills required to read, understand, and apply the explanations within ALEKS to solving problems (Figure 24 and Figure 25).

Table 8 Random Forest summaries for predicting MATH 101 Success from Registrar data and Assessment 0

```
                         Sample size: 398
           Frequency of class labels: 350, 48
                     Number of trees: 1000

Confusion matrix:

          predicted
  observed Pass Fail class.error
      Pass  349    1      0.0029
      Fail    3   45      0.0625

    Overall error rate: 1.01%
```

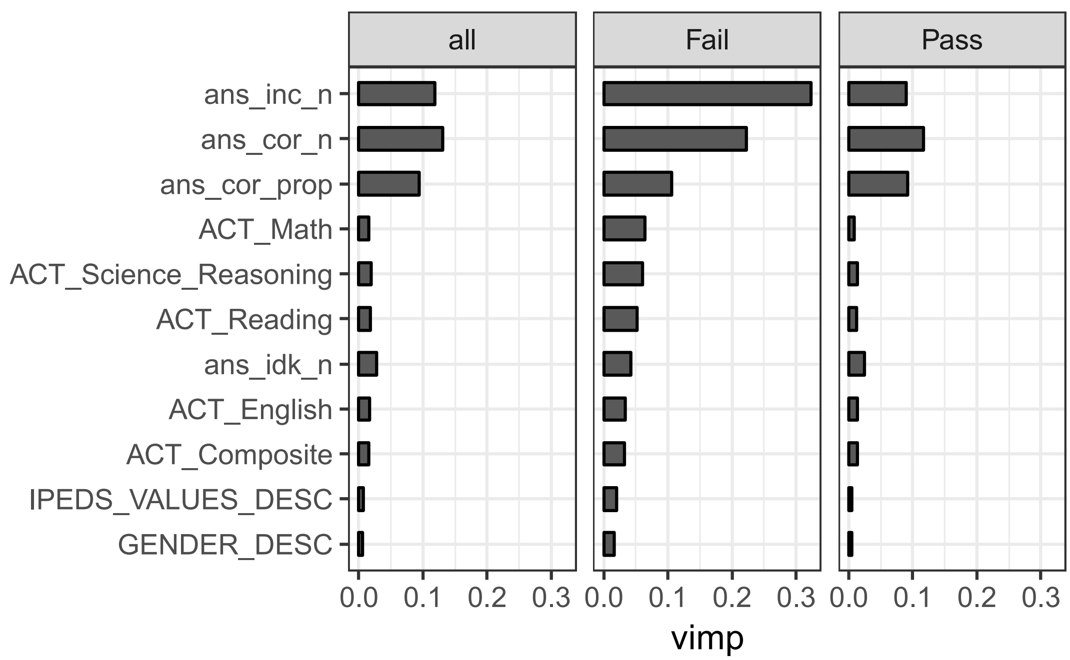

Figure 24 Random Forest VIMP for predicting MATH 101 Success from Registrar data and Assessment 0

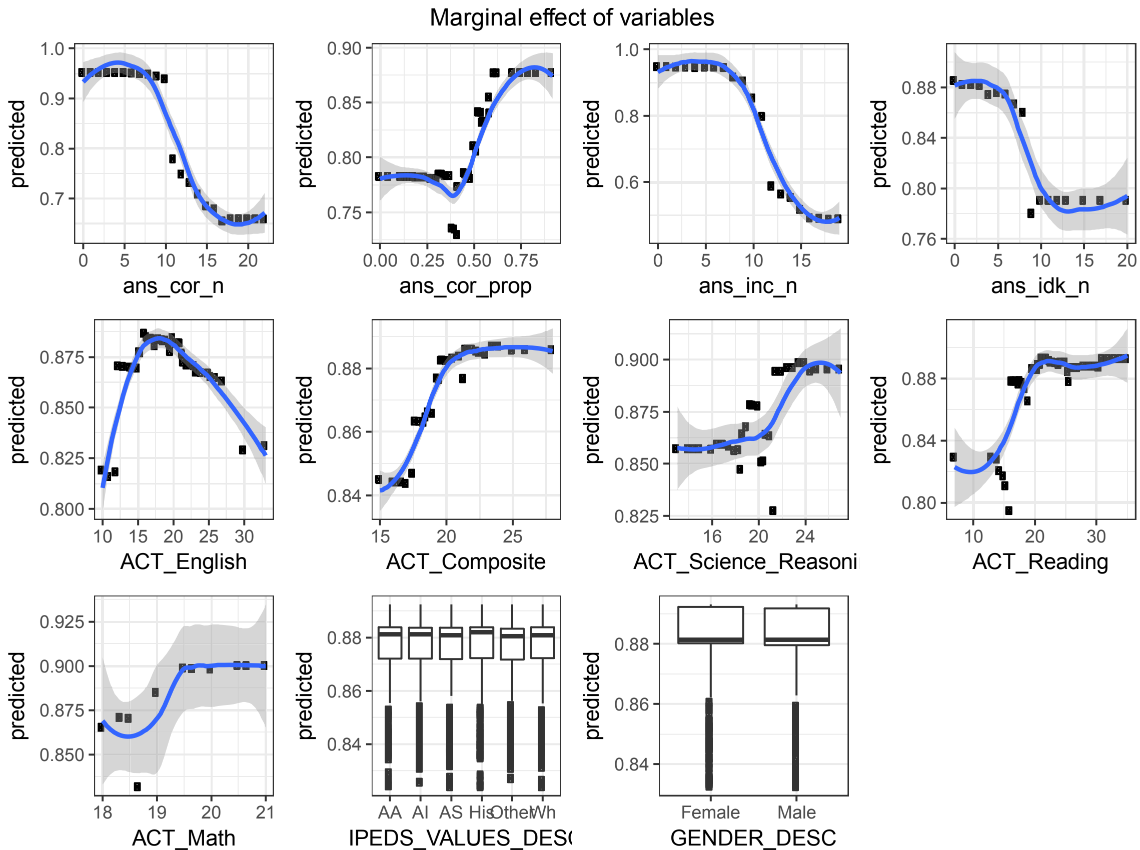

Figure 25 Random Forest Marginal Effects plot for predicting MATH 101 Success from Registrar data and Assessment 0

3.2.1.4 ACT Scores, demographics, and Practice question times predicting MATH 101 Success

The practice question time variables do not provide better prediction than the baseline assessment above. Results are similar to when Registrar data was not included (Figure 26 and Figure 27). The overall error rate of 2.46% is only slightly lower than without the Registrar variables above (Table 9).

Table 9 Random Forest summaries for predicting MATH 101 Success from Registrar data and Practice question timing

```
                     Sample size: 366
           Frequency of class labels: 319, 47
                     Number of trees: 1000

Confusion matrix:

          predicted
  observed Pass Fail class.error
      Pass  318    1      0.0031
      Fail    8   39      0.1702

    Overall error rate: 2.46%
```

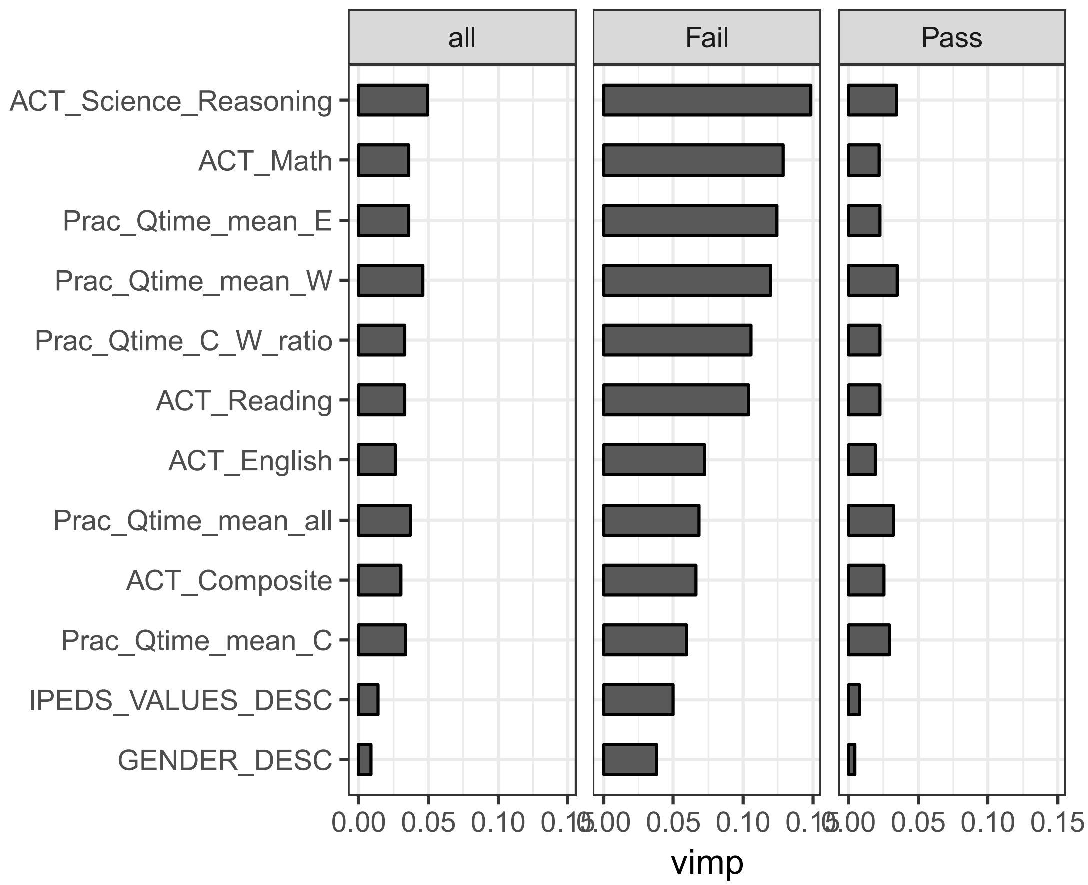

Figure 26 Random Forest VIMP for predicting MATH 101 Success from Registrar data and Practice question timing

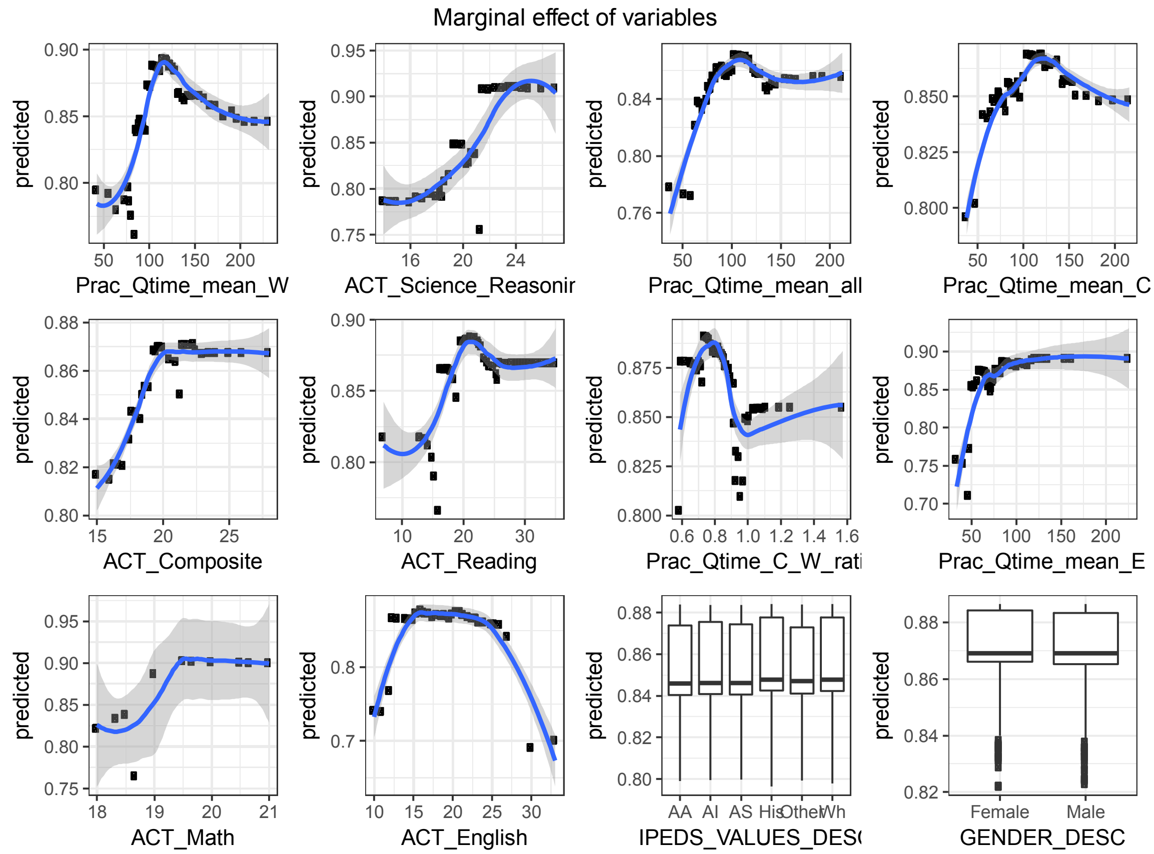

Figure 27 Random Forest Marginal Effect plot for predicting MATH 101 Success from Registrar data and Practice question timing

4 Discussion

Computer-based instruction has been successful. Just as others have seen increases (5%) in computer-based Intermediate Algebra courses versus face-to-face traditional lecture courses (Erin E. Krupa, 2014), UNM has experienced even larger increases with ALEKS in the emporium model (Figure 3), at least a 10% increase when combining MATH 101 and 102 compared to MATH 120 at Albuquerque/Main campus. The structure of the three 1-credit courses MATH 101-102-103 also improves success by forcing students to relearn topics if not mastered the first time. It may take a student more time on certain topics but the likelihood for overall student success is much higher with self-paced adaptive learning (Carlo A. Twigg, 2011). An additional side benefit of the emporium model has been that students can move along each semester up to 6 credit hours with MATH 101-102-103 and MATH 121 (College Algebra); students can save money when taking between 15-18 credit hours at UNM.

ALEKS is not the only innovation to help students. The MATH 101 instructors conduct experiments most semesters to further improve student success, and implement these on a rolling basis. The course modifications listed below come from discussions with Dr. Srini Vasan, Coordinator of Education Support for MATH 101, and from reviewing the syllabi over the last few years. One item to note is that students are encouraged to complete the course more quickly, resulting both higher success and earlier completion during the course (Figure 28).

The incremental increasing success rates each semester illustrates that continual improvement is possible with a course coordinator and staff who are actively engaged in meeting the needs of students for success. The ongoing support of The University, stakeholder Colleges, and Mathematics and Statistics Department have helped make this possible.

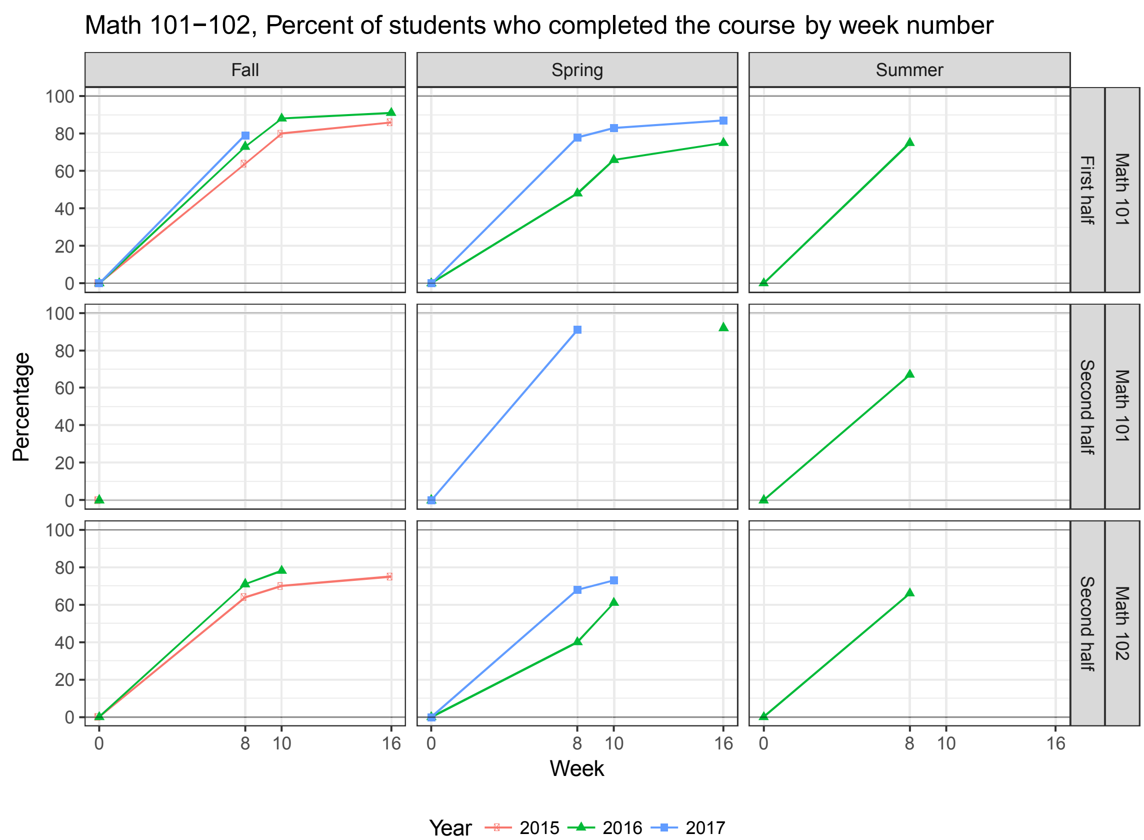

Figure 28 Student completion by week as summarized by Dr. Srini Vasan

The list below highlights the extensive incremental improvements incorporated into this course since moving to the emporium model.

Summer 2013-Spring 2014 (Ross, 2013)

The 3-credit 16-week (1-semester) Intermediate Algebra (MATH 120) course was split into the series of 1-credit 8-week MATH 101-102-103 individual courses.

- Student with three or more absences can be dropped by the instructor
- The Math Mαll Lab opens Monday to Friday from 8 AM to 6 PM and Sunday from 12 PM to 6 PM during semester long. During the interval break, The Math Mαll Lab opens Monday to Friday from 9 AM to 5 PM. Students can come during their free time for tutor or take exam. They only need to be to the Math Mαll Lab 2 hours before the place is closed for taking exam, otherwise, anytime for tutor.
- Each class has 5 tutors and 1 peer mentor tutor plus instructor and coordinator to support student's need during semester; during the interval break, the coordinator and tutors are available for student taking exam or need help with tutoring.
- The minimum requires that student must do 10 topics per week
- Required Benchmark is in week 3 and week 6
- Complete the pie and take exam latest by the end of week 8
- Student gets 2 attempts for the exam. The highest score will record for final grade of the course

- Letter grade for this class will be A, B, C or I, NC. Student must score 75% or above to be able to take the next class, otherwise, student will relearn the topics and retake the exam after fill up the pie again.
- The incomplete grade will turn to no-credit (NC) after 4 months if student does not complete the course

Fall 2014-Spring 2015 (Ross, 2014)

Required to come to the Math Mall Lab extra two hours/week if student is behind the required Benchmark to help student catch up and stay on track with the schedule.

- Workshop is held by peer mentor tutor around week 7 and week 14 or 15 depends on the semester and located at Union Building study room

Fall 2015-Spring 2016 (Vasan, 2015)

- ACT Math score was 18 instead of 19 is placed in MATH 101.
- The Math Mαll Lab website was developed with all resources to support student's need.
- Workshop is held at the Math Mαll Lab instead of someplace else in difference times and date during week 6 and 7 of the first half semester, and in week 14 and 15 in the second half of the to hope that more students come to try to complete the course and review for the exam.

- Pushing student complete the course by only given the break interval between semesters up until the Friday before the next semester begins to avoid getting no-credit instead of 4 months.
- Open new section for University of College by transfer student from Foundation Math, UNIV 103 with ACT math score 17 and below take MATH 101 if student pass the diagonal quiz with 75% and above. Also, take more students after first half semester to take MATH 101 from Foundation Math if they completed the course fast enough.
- Spring 2016, Facebook page was created to update with technology as student may use a lot , so hope that more students check it out with information about the workshop, times opening for the Math Mαll. Also video for some of topics based on during do tutoring and the exam result.

Fall 2016-Spring 2017 (Vasan, 2016a)

- Imply critical Benchmark by Thursday, 5 PM of week 5 to help student motivate to stay on the schedule for math 101 to be able to keep math 102 in the second half of the semester
- Critical Benchmark was 115 of 160 topics completed by the end of week 5 to encourage students are motivate to work harder to make the goal in the class. If student is below 115 topics, student is dropped from MATH 102 and sign up to UNIV 104 , Math Strategy to help student learn how to manage time, find resources around campus, getting if student is struggling, etc.

- Student still have a chance to get back to MATH 102 if student get back on track to the schedule and take exam by Friday, last day of the first half of semester. It is open up to the instructor to decide to take student back in.

Fall 2017 (Vasan, 2017)

- Required Benchmark was pushing up one week ahead of the schedule to help student complete the pie by the end of week 7 or week 15 instead of week 8 and week 16. So, student has time to wrap it up during weekend and take exam during week 8 or week 16.
- Critical Benchmark was pushing up to one week ahead of schedule (115/160 to 124/160 topics completed) by the end of week 5 to encourage students are motivate to work harder to make the goal in the class.
- Student with two or more absences can be dropped by the instructor.
- Student who has no access to ALEKS during the second week of classes may be dropped by the instructor.
- Any class period during which the student has no ALEKS access may be counted as an absence for the student.
- Anyone with no ALEKS access for 5 or more calendar days may be dropped from the class
- To prepare student for the next math courses, require student to write out the work in his/her notebooks. Bring the notebook to the lab every class. The instructor may perform periodic audits to ensure that student are

maintaining a proper notebook; failure of which could lead to additional assignments being given to the student.

- Practice exam was posted on the website, but then implemented in ALEKS in Fall 2017 (for same format) with multiple times to practice to help prepare students to take the exam
- Changing the exam (20 vs 28 questions), both to shorten length of time and questions for exam to increase success.
- Students are completing the course faster since they are being pushed harder.

The primary result found in this thesis is that by using the initial baseline assessment and student ACT scores (gender and race contributed no explanatory power) that we could predict whether a student was likely to succeed in MATH 101 with only a 1% error rate (Table 8). This could allow us within the first week of class to identify students who are most likely to fail and focus resources to help them succeed. The ACT English, Scientific Reasoning, and Reading component scores all contributed to student success, even more so that the Math component within the 18-21 range.

The ACT variables associated with English reading comprehension are predictive of student success in MATH 101. We hypothesize that this is because a student learns from ALEKS by reading and understanding the explanations before attempting to solve problems. It is possible that a student with weak language skills but average-to-strong math skills may still do poorly in a math placement test due strictly to their inability to understand written English. We may be seeing some students who place into UNIV 103 or MATH 101, not because of mathematics deficiencies, but because of English deficiencies. It may be worth considering, for a student with weak English skills, to defer taking their “remedial” mathematics courses until after they have improved their English skills. Such a student might then redo a math placement test and, with their newly improved English comprehension, may place much higher than they did previously. Thus, Math placement could be improved with a more comprehensive consideration of the many skills that are required for engaging with and learning mathematics.

5 Conclusions

At UNM, Intermediate Algebra success rates have greatly increased due to the implementation of the computer-based system ALEKS in the emporium Math Mαll, the conversion of MATH 120 to MATH 101-102-103, and the many incremental improvements made by the course coordinator. The statistical model developed in this thesis can be used on the first day of class to predict whether a student will pass with a 1% error rate to identify those students most likely to fail for the purposes of intervening to help those at-risk students succeed. This represents another potentially important incremental improvement to a series of successes in redesigning Intermediate Algebra at UNM.

5.1 Future work

We have several recommendations for further study. We would like to follow-up with students whom we predicted would do poorly and better understand what they struggled with. If there is a common pattern among these students, then that might suggest a type of intervention for these students who initially performed poorly on the baseline assessment and who then went on to fail the course. MATH 102 has a lower success rate than MATH 101, so a continued study of this course toward improving success is desired. Furthermore, the comparisons of 3-credit MATH 120 to 1-credit MATH 101 may not be entirely appropriate, MATH 102 should be included as part of a fairer comparison. The final exams have been changed to use a subset of the original 28 questions for a 20-question exam, and student success has improved; while fewer items are tested, we'd like to understand why students succeed with fewer items – perhaps less fatigue. Students now also have two attempts for final exams, with the second attempt focusing only on the questions answered incorrectly on the first attempt; is there a general pattern for the types of questions students get wrong on the exam? This would indicate insufficient preparation during the course for certain topics.

Regarding mathematics placement, students currently go into MATH 101 with an ACT Math score between 18 and 21. This semester (Fall 2017), if a student placed in UNIV 103 by having an ACT Math score no higher than 17 *and* the students scored well in the AcuPlacer they can go directly to MATH 101. Follow-up of these students will help us know whether the AcuPlacer is a good substitute for the initial ACT Math score for mathematics placement. Finally, consideration of the ACT English language skills might also help in placement or in the timing of MATH 101 over a student's career.

6 References

Adelman, C. (2006). The toolbox revisited: Paths to degree completion from high school through college. Washington, DC: U.S. Department of Education. Retrieved from https://www2.ed.gov/%0Arschstat/research/pubs/toolboxrevisit/toolbox.pdf

Attewell, P., Lavin, D., Domina, T., & Levey, T. (2006). New evidence on college remediation. *Journal of Higher Education, 77, 886–924.*

Bailey, T., Jeong, D. W., & Cho, S.-W. (2010). Referral, enrollment, and completion in developmental education sequences in community colleges.

Breiman, L. (2001a). Random Forests. Machine Learning, 45, 5-32. Retrieved from https://link.springer.com/article/10.1023%2FA%3A1010933404324?LI=true

Breiman, L. (2001b). Statistical modeling: The two cultures (with comments and a rejoinder by the author). Statistical science, 16(3), 199-231. Retrieved from https://projecteuclid.org/download/pdf_1/euclid.ss/1009213726

Bressoud, D. (2014). The pitfalls of pre-calculus. Retrieved from http://launchings.blogspot.com/2014/10/the-pitfalls-of-precalculus.html

Brown, R. S., & Niemi, D. N. (2007). Investigating the alignment of high school and community colleges. Retrieved from http://www.highereducation.org/reports/brown_niemi/BROWN_NIEMI.pdf

Bulger, S., & Watson, D. (2006). Broadening the definition of at-risk students. The Community College Enterprise.

Calcagno, J. C., & Long, B. T. (2008). *The impact of postsecondary remediation using a* *regression discontinuity approach: Addressing endogenous sorting and noncompliance*.

Christensen, R. (2006). *Log-linear models and logistic regression. Springer Science & Business Media.*

Cohen, A. M., Brawer, F. B., & Kisker, C. B. (2014). *The American community college* (6th ed.). John Wiley & Sons.

College, U. (2017). UNM Main Campus SAT/ New SAT/ ACT/COMPASS /ACCUPLACER Placement Table. Retrieved November 14, 2017, from http://ucollege.unm.edu/advisement/forms/2017-placement-table.pdf

Complete College America. (2012). Guided pathways to success: Boosting college completion. Retrieved from http://completecollege.org/docs/GPS_Summary_FINAL.pdf

Complete College America. (2016). Corequisite remediation: Spanning the completion divide. Retrieved from http://completecollege.org/spanningthedivide/

Cuseo, J. (2007). The empirical case against large class size: Adverse effects on the teaching, learning, and retention of first-year students. *The Journal of Faculty Development*, *21*(1), 5–21.

Erhardt, E., Bedrick, E., & Schrader, R. (2013). Lecture notes for {Advanced Data Analysis} 1 ({ADA1}), {Stat} 427/527, {University of New Mexico}},. Retrieved from http://statacumen.com/teach/ADA2/ADA2_notes.pdf

Erin E. Krupa, C. W. & J. M. (2014). *Undergraduate Students' Knowledge of Algebra:*

Evaluating the Impact of Computer-based and Traditional Learning Environments. Retrieved from 2014_Krupa_UndergraduateStudentsKnowledgeofAlgebraEvaluatingtheImpactofComputerbasedandTraditionalLearningEnvironments.pdf

Falmagne, J.-C., Cosyn, E., Doignon, J.-P., & Thiéry, N. (2006). The assessment of knowledge, in theory and in practice. In *Formal concept analysis* (pp. 61–79). Springer.

Ishwaran, H., & Malley, J. D. (2014). Ishwaran, H., & Malley, J. D. (2014). Synthetic learning machines. BioData mining, 7(1), 28.

Knowlton, E. N. (2011). An Intervention Strategy Designed to Reduce Math Remediation Rates at Community Colleges. Retrieved from TandFonline

Massy, W., & Zemsky, R. (1995). *Using information technology to enhance academic productivity*.

McGraw Hill Education. (2017a). Overview of ALEKS. Retrieved from https://www.aleks.com/about_aleks/overview#origine

McGraw Hill Education. (2017b). Success Stories. Retrieved from https://www.aleks.com/about_aleks/success_stories

McGraw Hill Education. (2017c). What is ALEKS ? Retrieved from https://www.aleks.com/about_aleks

Mehaffy, G. L. (2012). Challenge and Chance. Retrieved from http://www7.nau.edu/plcc/Reading/Challenge and Change Article.pdf

Rankin, S. (2016). Developing College Foundations through Foundational Math. Retrieved from http://www.hed.state.nm.us/uploads/files/Policy and Programs/Articulation and Transfer Reform/NMHED Math Remediation.pdf

Ross, J. (2013). MATH 101 Syllabus.

Ross, J. (2014). MATH 101 Syllabus.

Steve Carr. (2013). Grand Opening Set for Math Learning Lab at UNM. Retrieved from http://news.unm.edu/news/grand-opening-set-for-math-learning-lab-at-unm

Suilmann, J. (2015). UNM discontinues remedial Introductory Studies courses. Retrieved from https://news.unm.edu/news/unm-discontinues-remedial-introductory-studies-courses

The University of New Mexico. (2016). The Department of Mathematics and Statistics, MATH 101, 102, 103. Retrieved November 14, 2017, from http://math.unm.edu/courses/materials/math-101-102-103

The University of New Mexico. (2017). College Enrichment Program. Retrieved November 14, 2017, from cep.unm.edu/academic-support/success-navigator/index.html

The University of New Mexico. (2017). Office of Institutional Analytics. Retrieved November 14, 2017, from http://oia.unm.edu/facts-and-figures/official-enrollment-reports.html

The University of New Mexico. (2017). Office of Institutional Analytics, Academic Program Review. Retrieved November 14, 2017, from http://oia.unm.edu/facts-and-figures/apr data.html

The University of New Mexico. (2017). University Catalog, Mathematics and Statistics. Retrieved November 14, 2017, from http://catalog.unm.edu/catalogs/2017-2018/colleges/arts-sciences/math-stat/index.html

The University of New Mexico. (2017). University Catalog, The Undergraduate Program. Retrieved November 14, 2017, from http://catalog.unm.edu/catalogs/2017-2018/undergrad-program.html

The University of New Mexico. (2017). UNM Bureau of Business and Economic Research. Retrieved November 14, 2017, from http://bber.unm.edu/county-profiles

Twigg, C. A. (2011). The math emporium: A silver bullet for higher education. *Change: The Magazine of Higher Learning*, *43*(3), 25–34.

Twigg, C. A. (2011). The math emporium: Higher education's silver bullet. Retrieved from https://www.researchgate.net/publication/254338451_The_Math_Emporium_A_Silver_Bullet_for_Higher_Education

United States Census Bureau. (2010a). American Fact Finder. Retrieved November 14, 2017, from factfinder.census.gov/bkmk/table/1.0/en/ACS/14_5YR/DP04/0400000US35

United States Census Bureau. (2010b). *County in each state in America*. Retrieved from https://www2.census.gov/geo/docs/reference/codes/files/national_county.txt

Vasan, S. (2015). MATH 101 Syllabus.

Vasan, S. (2016a). MATH 101 Syllabus.

Vasan, S. (2016b). Math Learning Lab (MaLL) Home Page. Retrieved from http://math.unm.edu/~svasan01/

Vasan, S. (2016c). Personal Interview.

Vasan, S. (2017). MATH 101 Syllabus.

Wang, X. (2014). *Bayesian Partially Ordered Probit and Logit Models with an Application to Course Redesign*.

Ward, L. (2017). Working using ALEKS in Intermediate Algebra since 2012.

Wickham, H. (2009). ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York.

Wimberly, G. L., & Noeth, R. J. (2005). College Readiness Begins in Middle School. ACT Policy Report. *American College Testing ACT Inc*.