Modeling the Effectiveness of Developmental Instruction

Yelena Perkhounkova

Julie Noble

Richard Sawyer

ACT

August 2005

Modeling the Effectiveness of Developmental Instruction

Yelena Perkhounkova
Julie Noble
Richard Sawyer

Abstract

Debates about developmental instruction have prompted postsecondary institutions to justify their developmental courses. This study investigated a method of documenting the benefits of developmental courses that involves using placement test scores to predict college outcomes separately for students who take a developmental course before taking standard courses and for those who enroll in standard courses directly. College outcomes were defined in terms of success in standard courses or retention in college. If, among students with a given placement test score, the predicted probability of success or of retention is higher for students who took a developmental course, one would have evidence that this developmental course is beneficial. Multi-year data from two large Midwestern universities were used to illustrate this approach. Simply taking a developmental course did not usually result in benefits. Students who earned high grades (at least a B) in developmental courses, however, were more likely than other students to succeed in standard courses or to persist in college.

Modeling the Effectiveness of Developmental Instruction

Postsecondary institutions offer developmental courses to students who are academically underprepared so that they can acquire the knowledge and skills required to succeed in standard college-level courses. According to McCabe (2000), 41% of community college enrollees and 29% of four-year college enrollees are inadequately prepared in at least one of the basic skills areas (writing, reading, and mathematics). As a result, close to 30% of U.S. freshmen enroll in developmental programs as they enter postsecondary education (National Center for Education Statistics, 2000). In 2001-02, developmental instruction was provided by nearly all public two-year postsecondary schools, 80% of public four-year institutions, and approximately 65% of private four-year institutions (NCES, 2003). Courses whose purpose is to prepare students for standard-level college courses are designated as "remedial," "developmental," or "compensatory," depending on the institution. In this report, we use the term "developmental."

American postsecondary institutions have a long tradition of providing developmental education. Recently, national debate about teaching college students the skills that many think they should have learned in high school has prompted scrutiny of the costs and benefits of developmental education. While some are concerned with the cost of providing developmental instruction (estimated at $1 billion, or 1% of all public expenditures for postsecondary education in the United States, according to Phipps, 1998), others consider this amount a worthy investment that benefits society and the institutions that provide developmental services (Phipps, 1998; Spann, 2000).

Regardless of financial concerns, a central question in this debate is whether developmental instruction provides educational benefits to students. To date, research on this issue has been inconclusive. For example, Weissman, Silk, and Bulakowski (1995) found that

academically underprepared students who took developmental courses performed about as well as average college-level students, while other underprepared students who took no developmental courses attempted and earned fewer credits and had significantly lower cumulative GPAs (see also Schoenecker, Bollman, & Evens, 1996; Chen & Cheng, 1999). On the other hand, a national longitudinal study of 1982 high school graduates concluded that full-time entering freshmen who enrolled in developmental courses continued at their institution to the start of their second year at a somewhat lower rate than all full-time entering freshmen (NCES, 1996). Moreover, according to the same study, students who took one or more developmental courses had lower rates of bachelor's degree completion, compared to students who took no developmental courses. In part due to these contradictory findings, institutions and states have been pressured by policy makers and the public to justify their developmental courses.

According to Phipps, an institution that provides developmental instruction should address the following questions to evaluate its effectiveness: (1) Do students successfully complete developmental courses? (2) Do students move from developmental instruction to college-level work? (3) Do students who take developmental courses eventually complete college-level courses? (4) Are developmental instruction students persisting and reaching their academic goals? Various methods have been employed to evaluate the effectiveness of developmental instruction, including tracking student progress (Weissman et al., 1995), modeling the need for remediation (Schoenecker et al., 1996), and posttesting (Sawyer & Schiel, 2000). One method of documenting developmental course benefits is to advance a "value-added" argument: Do underprepared students benefit at all from taking developmental courses? And if

they do, by how much? Two general approaches to demonstrating added value are posttesting and follow-up tracking.

Posttesting involves administering a placement test to students at the beginning and at the end of developmental instruction. If a developmental course is effective in teaching students the required skills, then students' scores on the placement test at the end of the course should be higher than their scores at the beginning of the course. Ideally, they should meet the cutoff established for enrollment in the standard course.

Follow-up tracking requires collecting data on students' academic performance after they complete developmental course work. The relationship between the measure used for course placement and outcome criteria is documented separately for students who completed developmental instruction and for those who did not. Sawyer (1997) proposed using logistic regression analysis to model the relationship between placement test scores and academic performance for two groups of students—those who took a developmental course before taking standard courses and those who enrolled in standard courses directly. Effectiveness of a developmental course would be demonstrated when the probability of success for students with a given initial test score who took a developmental course exceeds the probability of success for students with the same score who did not take the developmental course.

The objective of this study was to investigate the follow-up tracking method of documenting developmental course benefits. We considered two kinds of academic performance outcomes—success in particular standard courses and retention in college. The predictor variables were ACT scores and ACT scores and developmental course grades jointly. We also applied the same methodology to pairs of courses for which one was at a lower level than the other; neither course had to be specifically designated as "developmental" by the institution.

Method and Data Sources

This study used data from two large Midwestern universities (referred to as Institution I and Institution II), and the data sets for the two institutions were analyzed separately. Both institutions determined placement in developmental mathematics and English courses by a combination of local placement test scores and appropriate ACT scores (English or Mathematics). Other variables, such as available transfer credit, were also considered in placement decisions. For practical reasons, models investigated in this study did not include placement variables other than ACT scores. Although such models represent a simplified view of actual placement practices, they are useful in evaluating the effectiveness of developmental courses.

To ensure adequate sample sizes by course, data were pooled from several entering classes at each institution. Institution I data consisted of the ACT test scores and college course grades of the 1997 through 2002 entering classes. Institution II data consisted of the ACT test scores and college course grades of the 1996 through 2000 entering classes.

Logistic Regression Models

Logistic regression is often used to model the statistical relationship between test scores and outcome criteria coded as 0 (failure) or 1 (success). We developed separate logistic regression models for developmental students and for students who enrolled in standard courses directly. The models predict either students' conditional probability of success in standard courses or their conditional probability of retention:

$$\hat{p} = \frac{1}{1 + e^{(-\text{Index})}}$$

where $\text{Index} = a_0 + a_1 x_1 + \ldots + a_n x_n$, a weighted combination of predictor variables $x_1, \ldots, x_n$. The regression coefficients $a_0, a_1, \ldots, a_n$ were estimated from data. We developed two types of

regression models, using as predictor variables either ACT test scores alone or ACT test scores and developmental course grades jointly.

Success in Upper-Level Courses

We analyzed data for different pairs of lower-level and upper-level courses, which are listed in Table 1 along with their respective sample sizes. In general, the "lower-level" courses were developmental courses and the "upper-level" courses were standard courses. However, in some instances the "upper-level" courses were also developmental. The sample size (n) listed for each lower-level course is the number of students who took it and then took the upper-level course. The sample size listed for each upper-level course is the number of students who enrolled in it directly. Records of students who did not enroll in the upper-level course after taking a lower-level course were excluded from the analysis. Because Institution II data contained more detailed information than Institution I data, we were able to select for our developmental samples only those students who took no other courses in the relevant subject area between paired lower-level and upper-level courses. For example, 67 students at Institution II who took Basic Algebra 2 and Trigonometry did not take any other mathematics courses in the time period between these two courses. This cannot be stated with certainty for the developmental samples from Institution I.

TABLE 1

Course Pairs and Sample Sizes

Lower-level course	n	Upper-level course	n
Institution I			
Developmental Math 1	81	Developmental Math 2	681
Developmental Math 2	385	Introduction to Algebra	1,330
Introduction to Algebra	529	Algebra and Pre-Calculus	1,136
Basic Writing 1	130	Basic Writing 2	1,399
Basic Writing 2	772	Composition and Rhetoric	3,096
Institution II			
Basic Algebra 2	67	Trigonometry	456
Basic Algebra 2	365	Calculus	2,979
Rhetoric 1	6,215	Rhetoric 2	725
Rhetoric 2	3,609	Literature	2,456
Accelerated Rhetoric	3,650	Literature	2,456
Rhetoric W	217	Literature	2,456
Rhetoric S	492	Literature	2,456
Rhetoric 2	178	Creative Writing	241
Accelerated Rhetoric	163	Creative Writing	241
Rhetoric S	77	Creative Writing	241

In the first stage of the analysis, logistic regression models were developed for predicting the conditional probability of success in the upper-level courses separately for students who took the lower-level courses and for those who did not. Success was defined in two ways: as a B or higher, or a C or higher, course grade. Confidence bands (95%) were constructed around the probability curves to identify statistically significant differences.

In the next stage of the analysis, students' grades in lower-level courses were included with ACT scores for predicting the conditional probability of success in the upper-level courses. This allowed separate probability curves for students earning a D, a C, a B, or an A grade in the lower-level course. Course grades were originally coded on a 13-point scale (A+ through F). For this analysis, we recoded them to a traditional A–F five-point scale to conform to ACT Course

Placement Service format and to simplify the interpretation of the probability curves without losing essential information.

Retention

Next, we examined the relationship between developmental instruction and retention. Logistic regression models were estimated for the three criterion variables of interest: enrollment in the second term (Institutions I and II), enrollment in the first fall term following initial enrollment (Institutions I and II), and enrollment in the second fall term following initial enrollment (Institution II). Separate models were estimated for students who received developmental instruction and for those who did not. ACT score alone and ACT score and developmental course grade jointly were used as predictors.

Sample sizes for this analysis are listed in Table 2. The samples were limited to students who had enrollment information for the terms investigated in this study. The developmental samples included students who took a particular developmental course, and the non-developmental samples included students who did not take any developmental courses in the corresponding subject area.

The Institution I data included enrollment status indicators for the second term and first fall term following initial enrollment. Of this institution's 8,787 students, 593 did not enroll in the second term after initial enrollment and 1,004 did not enroll in the first fall term after initial enrollment. Students who first enrolled in fall 2002 (n = 2,528), the final year of the study, were not included in the models for one-year retention because no information was available about their second-year enrollment at the time of data collection. Because Institution I data included no information about the terms under which particular courses were taken, a small percentage of students in the developmental samples might have taken the developmental course during or

after the enrollment term under investigation. However, we have no reason to suspect that this possibility would have noticeably affected the results.

Although the Institution II data set did not include enrollment status indicators, it contained term information for all courses taken by each student. We used this information to construct enrollment indicators for the three terms of interest. This allowed inclusion in the developmental sample of only those students who took a particular developmental course prior to the enrollment term under investigation. Also, because the Institution II data were more detailed, we were able to select for the non-developmental samples only those students who took at least one standard course in the relevant subject area (mathematics or English).

TABLE 2

**Sample Sizes for Retention Analyses, by Course, for Developmental and
Non-Developmental Samples**

Developmental course	Enrollment	Developmental n	Non-Developmental n
Institution I			
Developmental Math 1, Developmental Math 2, or Introduction to Algebra	Second-term	2,179	6,608
	Second-year	1,619	4,640
Developmental Math 1	Second-term	201	6,608
	Second-year	163	4,640
Developmental Math 2	Second-term	762	6,608
	Second-year	586	4,640
Introduction to Algebra	Second-term	1,715	6,608
	Second-year	1,368	4,640
Basic Writing 1 or Basic Writing 2	Second-term	1,545	7,242
	Second-year	955	5,304
Basic Writing 1	Second-term	146	7,242
	Second-year	139	5,304
Basic Writing 2	Second-term	1,529	7,242
	Second-year	946	5,304
Institution II			
Basic Algebra 1 or Basic Algebra 2	Second-term	974	8,245
	Second-year	783	6,475
	Third-year	535	4,417
Basic Algebra 2	Second-term	896	8,245
	Second-year	721	6,475
	Third-year	495	4,417
Rhetoric 1, Rhetoric 2, Accelerated Rhetoric, Rhetoric W, or Rhetoric S	Second-term	13,004	1,144
	Second-year	10,711	1,055
	Third-year	7,016	945
Rhetoric 2	Second-term	443	1,144
	Second-year	4,879	1,055
	Third-year	3,203	945
Accelerated Rhetoric	Second-term	4,684	1,144
	Second-year	4,155	1,055
	Third-year	2,793	945
Rhetoric W	Second-term	161	1,144
	Second-year	137	1,055
	Third-year	94	945
	Second-term	1,044	1,144
	Second-year	1,034	1,055

| Rhetoric S | Third-year | 665 | 945 |

Results

Success in Upper-Level Courses

For each pair of lower-level and upper-level courses, we created graphs showing the relationship between students' ACT scores and their probability of success in the upper-level course. Each graph included a curve for students who took the lower-level course prior to taking the upper-level course and a curve for students who took only the upper-level course. If developmental instruction is effective, at least in the range of lower test scores, the probability-of-success curve for students who took the lower-level courses should lie above the probability-of-success curve for students who enrolled directly in the upper-level course. For Institution I, this expected outcome occurred for about 50% of the English course analyses and about 60% of the mathematics course analyses. For Institution II, this outcome occurred for only about 40% of the English course analyses and about 50% of the mathematics course analyses. Figure 1 illustrates the results for one such course pair. For this example, nearly all students (97%) who took Developmental Math 1 (dashed lines) had ACT Mathematics scores between 11 and 20. Compared to students in the same score range who did not take Developmental Math 1 (solid line), those who took it had a higher probability of earning at least a C in Developmental Math 2.

Figure 1 also includes 95% confidence bands around the probability-of-success curves, as shown by the lighter (dashed and solid) lines. For this example as well as for other course pairs, the confidence bands overlapped in the range of lower test scores, which indicates that the observed benefit from developmental instruction was not statistically significant at the .05 level. This outcome is partly explained by the small sample sizes in the lower part of the score range for many of the course pairs. For instance, the probability curves shown in Figure 1 are based on

the samples of 81 students who took Developmental Math 1 and 681 students who did not take it. The small sample sizes occurred despite our pooling data over several years.

FIGURE 1. Effectiveness of Developmental Math 1 in Improving Success in Developmental Math 2. (Success Criterion = C or Higher Grade in Developmental Math 2)
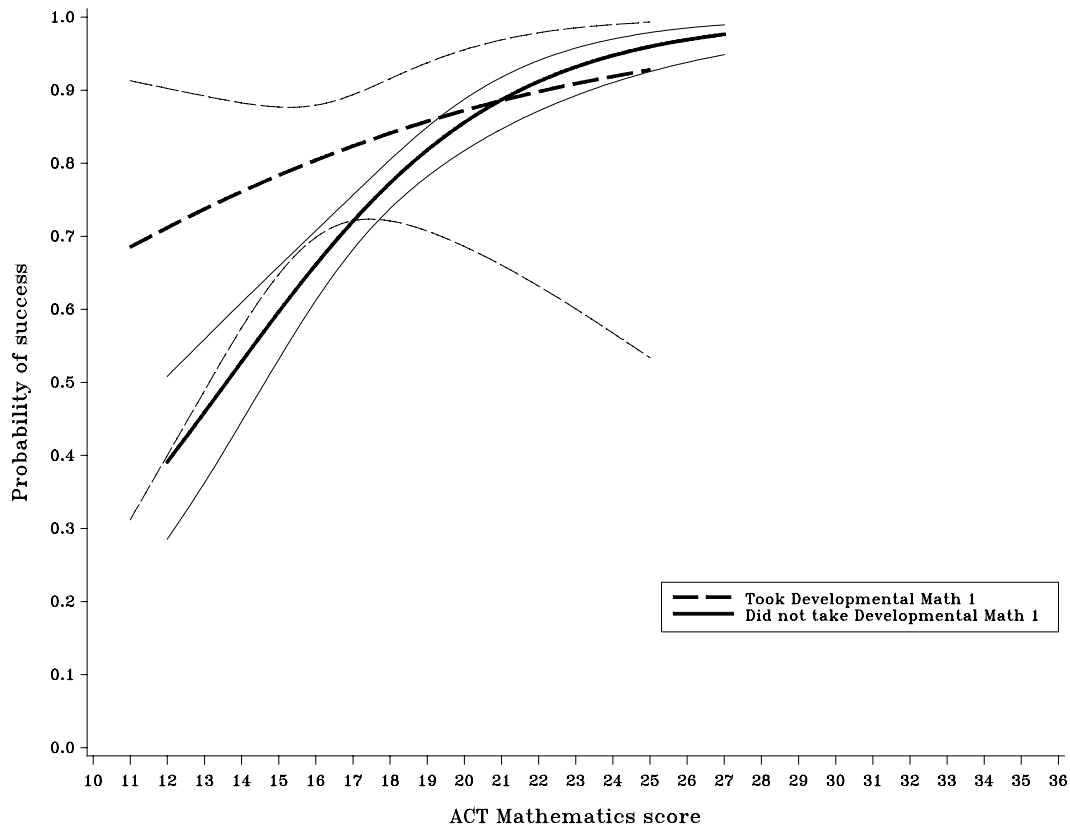

For Institution II, we were also able to compare the effects of several prerequisite rhetoric courses on success in two standard courses (Literature and Creative Writing). Most undergraduates at Institution II must satisfy an English requirement by taking one of four basic rhetoric courses (Rhetoric 1-Rhetoric 2 sequence, Accelerated Rhetoric, Rhetoric W, or Rhetoric S). Satisfactory completion of the rhetoric requirement is a prerequisite for taking Literature, a required course in the general education program. Rhetoric 1-Rhetoric 2, a two-semester course

sequence, and its accelerated version, Accelerated Rhetoric, cover the same content areas (speaking, writing, and reading). Rhetoric W emphasizes reading and writing and was designed for students who have fulfilled the speaking requirement. Rhetoric S emphasizes reading and speaking and was designed for students who have fulfilled the writing requirement. Students are assigned to these courses on the basis of their ACT English test scores and any transfer credit.

The results showed that taking Rhetoric 2 significantly increased the probability of success in the Literature course for students in the middle part of the score range (ACT score range of about 19 to 26) and taking Rhetoric S significantly increased the probability of success in the Literature course for higher-scoring students (ACT score range of 27 to 36). For illustration of this second course pair, see Figure 2. The majority of students (80%) who took Rhetoric S had ACT English scores in the range of 21 to 36. In this score range, the probability-of-success curve associated with taking Rhetoric S was higher than that associated with not taking any of the rhetoric courses. It appears that higher-scoring students (who probably had some writing transfer credit) were assigned to Rhetoric S. Other rhetoric courses had no statistically significant effects on success in the standard courses.

FIGURE 2. Effectiveness of Rhetoric S in Improving Success in Literature.
(Success Criterion = B or Higher Grade in Literature)
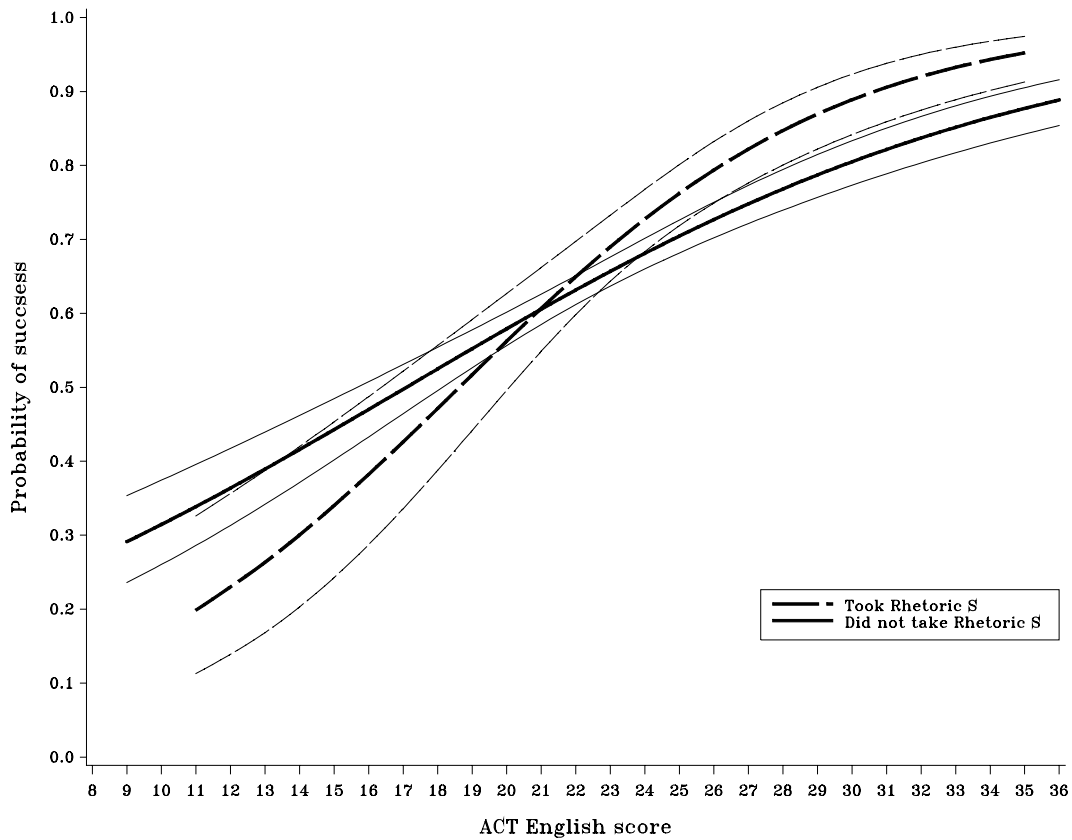
Another interesting finding was that for both groups of students (developmental and non-developmental) and both institutions, probability-of-success curves for English courses were generally higher and flatter than the probability-of-success curves for mathematics courses. This finding is consistent with other ACT course placement research (ACT, 2003).

When students' grades in lower-level courses were added to ACT scores for predicting the probability of success in upper-level courses, the results almost always indicated that earning a high grade (A or B) in a lower-level course increased students' probabilities of success in the upper-level course over enrolling in it directly. Students with C or D grades typically had lower

probabilities of success in the upper-level course than did students who enrolled in it directly. An example is shown in Figure 3.

FIGURE 3. Effectiveness of Basic Algebra 2 in Improving Success in Trigonometry, by Grade in Basic Algebra 2. (Success Criterion = B or Higher Grade in Trigonometry)
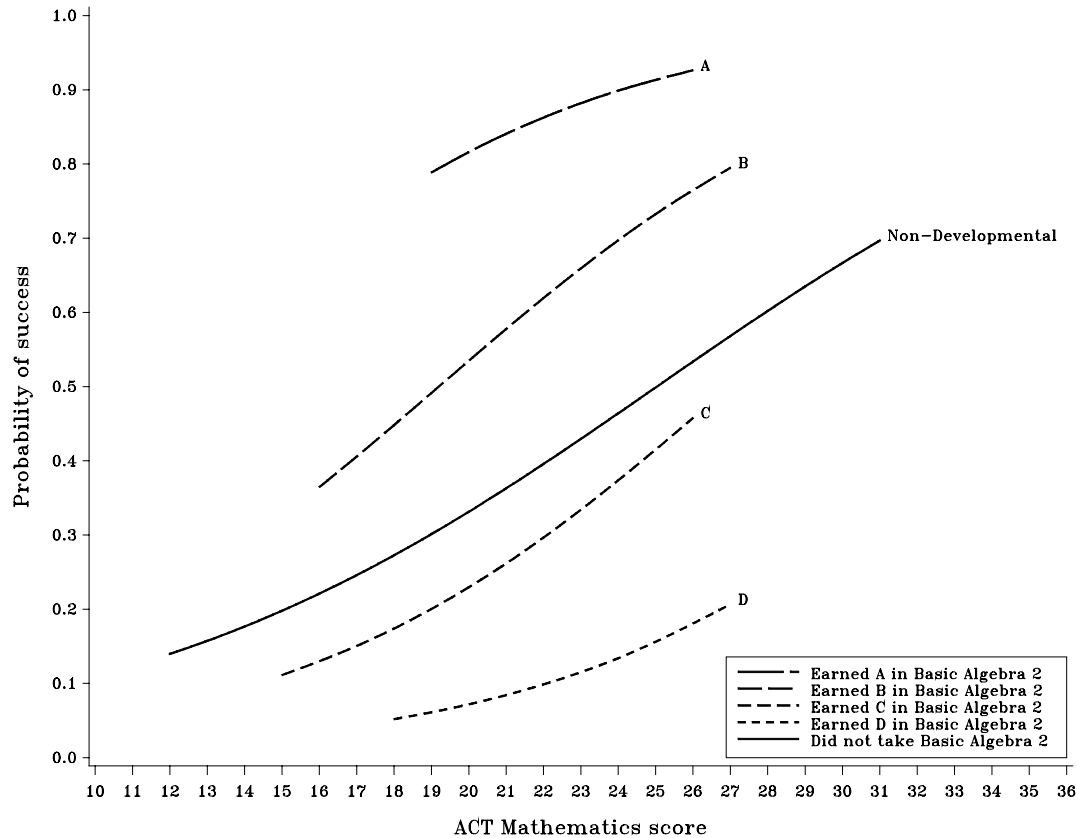

Retention

We created graphs showing the relationship between students' probability of retention and their ACT test scores. Each graph included two curves: one for students who took a developmental course and the other for students who took no developmental courses in the relevant subject area. For Institution I, the probability of retention curve (using both enrollment in the second term and enrollment in the first fall term following initial enrollment) for students who took a developmental course lay above the corresponding curve for students who took no

developmental courses, at least for students with lower ACT test scores. However, because the retention rate was already high (typically 80% or higher), the improvement in probability was often small. Generally, when retention is very high it is difficult to demonstrate substantive benefit.

Figure 4 shows an example of an effective developmental mathematics instruction program at Institution I. In the range of lower ACT test scores, the probability-of-success curve for students who took at least one of three possible developmental mathematics courses was higher than that for non-developmental students. Therefore, developmental mathematics students had a higher probability of second-year reenrollment than students who took no developmental mathematics courses. Eighty percent of developmental students had ACT Mathematics scores in the range of 11 to 21.

An unexpected result was that the relationship between the probability of retention and test scores was often near zero or even negative among students who took a developmental course (see Figure 4). This result suggests that students with higher ACT test scores who were given developmental instruction were less likely to reenroll than students with lower test scores.

FIGURE 4. Effectiveness of Developmental Mathematics Instruction at Institution I
(Developmental Math 1, Developmental Math 2 or Introduction to Algebra) in Improving
Second-Year Retention
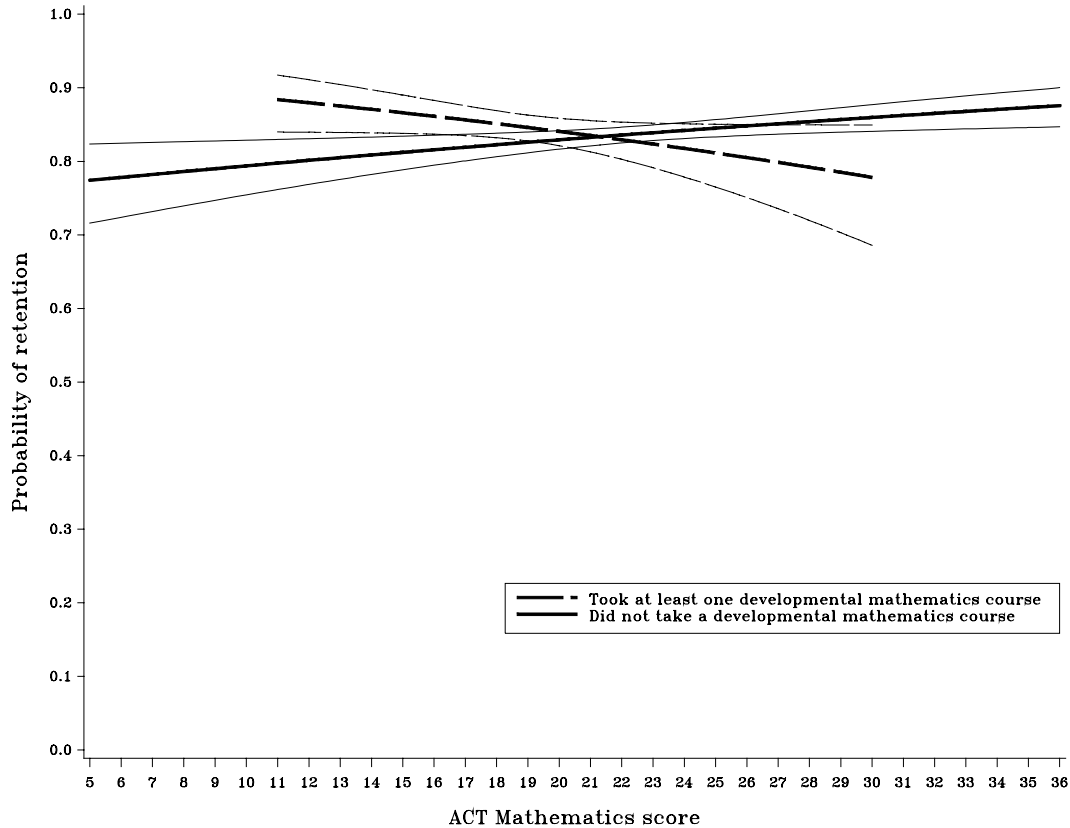
In contrast to the results for Institution I, most of the results for Institution II showed a

lower probability curve for students who took developmental courses, especially in the range of

lower ACT test scores. For example, Figure 5 shows that in the range of lower ACT test scores

students who took one or both of the developmental mathematics courses (Basic Algebra 1

and/or Basic Algebra 2) had a lower probability of second-year reenrollment than did students

who did not take any developmental mathematics courses. The probability-of-success curve was

lower for developmental students than for non-developmental students in the range of ACT

Mathematics scores from 12 to 24, where 88% of Institution II developmental students scored.

FIGURE 5. The Effect of Developmental Mathematics Instruction at Institution II (Basic Algebra 1 and/or Basic Algebra 2) on Second-Year Retention
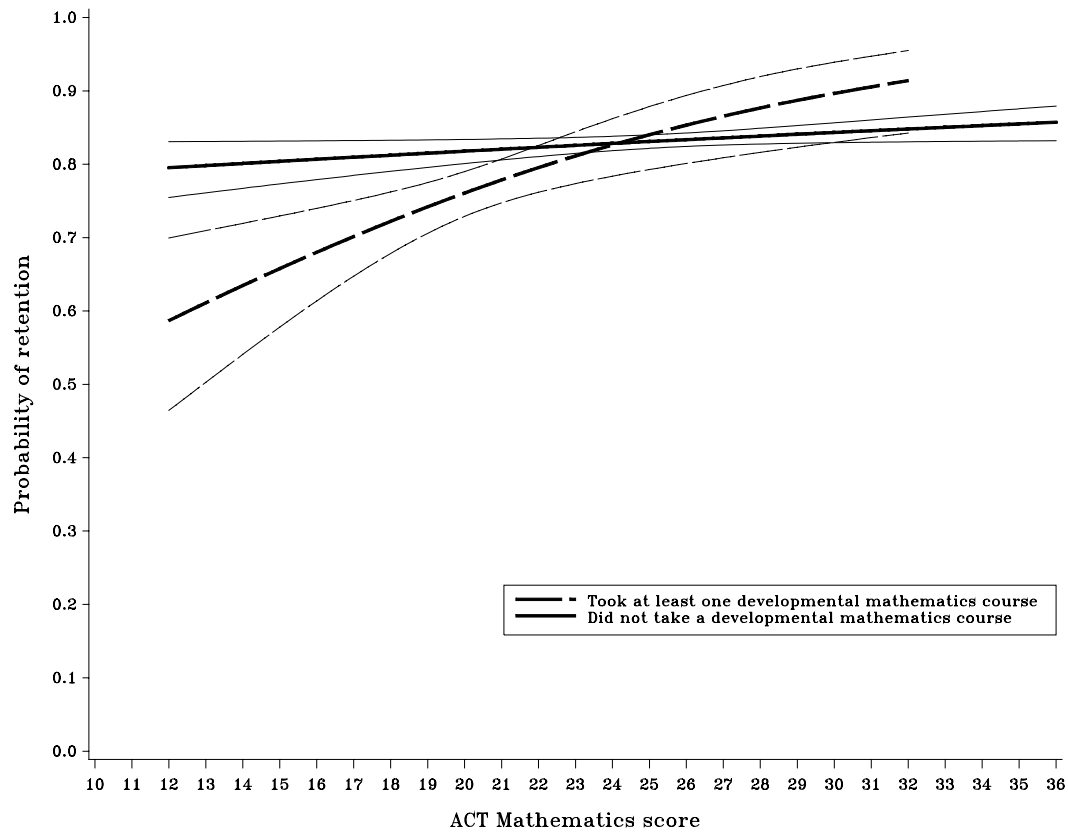
Next, we added students' grades in developmental courses to the models predicting students' probability of retention. Similar to the results of the course grade analyses, these results showed that students with lower ACT scores, and often higher-scoring students as well, who earned a higher grade (A or B) in a developmental course had higher probabilities of retention than those who earned a lower grade. Figure 6 illustrates the effect of students' grades in a developmental mathematics course (Basic Algebra 1 or Basic Algebra 2) on their probability of second-year reenrollment. Collectively, lower-scoring students who took a developmental mathematics course at the Institution II had a lower probability of registering for the second year than students with similar scores who enrolled directly in standard courses (as illustrated in

Figure 5). However, when students' grades in the developmental courses were taken into account, it became apparent that students who earned a B or A grade in the developmental mathematics courses had a higher probability of registering for the second year than did students who enrolled directly in standard mathematics courses. The probabilities of continuing to the second year at this school for students who did not take a developmental mathematics course were at least .15 lower than the corresponding probabilities for students who earned an A in their last developmental mathematics course (see Figure 6).

FIGURE 6. Effectiveness of Developmental Instruction at Institution II (Basic Algebra 1 and/or Basic Algebra 2) in Improving Second-Year Retention, by Developmental Course Grade.
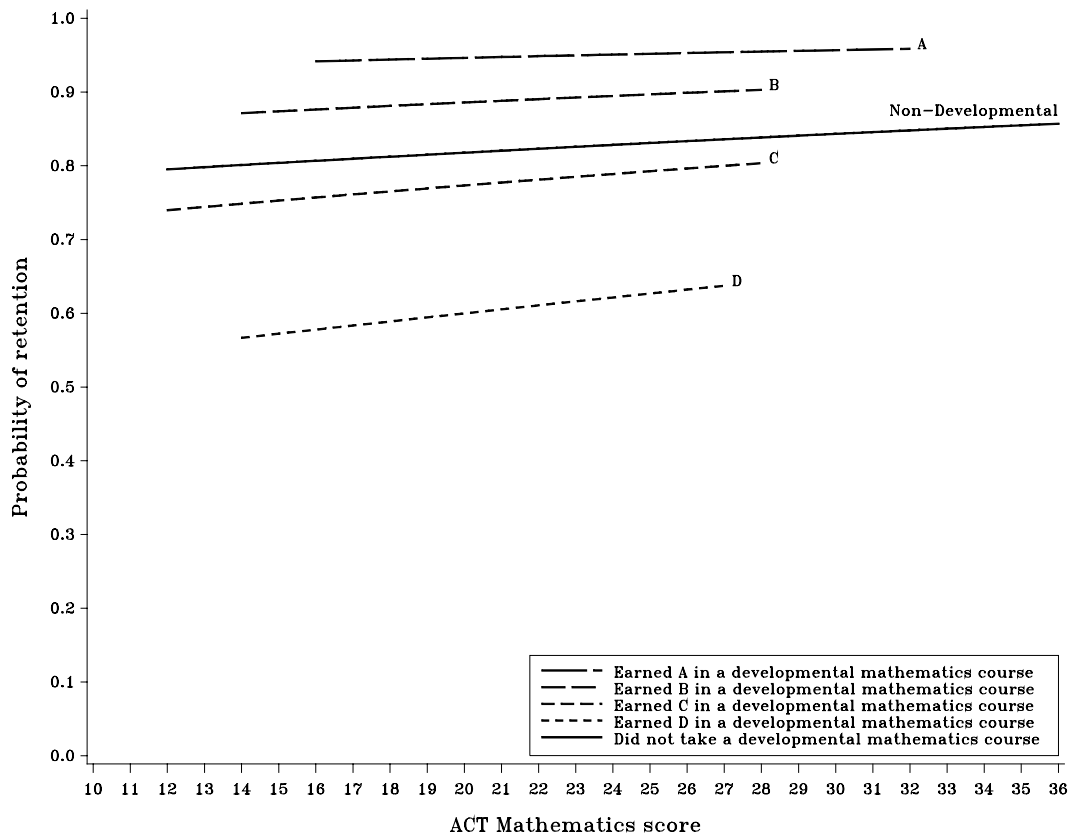
Discussion

Using data from two institutions, we showed that students who took developmental courses and earned high grades (at least a B) in these courses were more likely to be successful than were other students with similar levels of academic preparation. Success was defined as retention in college and grades in a subsequent higher-level course.

This study differs from previous approaches because it takes developmental course grades into consideration. The results show that simply comparing all developmental students to non-developmental students might lead to erroneous conclusions about the effectiveness of a developmental course. The *grade* earned by the student in a developmental course is a more important predictor of later success than merely taking the course. This result was true whether success was defined by the grade in a subsequent standard course or by retention. Although data provided by these institutions differed in terms of detail, the same general results were found for both institutions. A key limitation of this study was that it involved only two institutions. These analyses need to be replicated at more institutions to determine if the findings also hold true for them.

Another limitation of this approach is that finding statistically significant benefits may be hindered by small sample sizes, especially at lower score values. In general, the data for this type of analysis needs to contain information for many different courses taken over several years, as well as enrollment information for many students. Because the classes may not be large enough, information from several entering classes may need to be aggregated. Another possibility is to develop hierarchical models, which use data from several units (i.e., institutions) collaterally to improve the prediction of the models for each individual unit.

The main advantage of this study—that it was based on actual data from postsecondary institutions that provide developmental education—was also a limitation. Ideally, the best way to determine effectiveness of developmental instruction would be to randomly assign students either to the developmental course or to the standard course and then examine their later academic success. However, this approach would generally not be an option in most educational environments. We believe that this limitation did not affect the main findings of this study.

Documenting and managing information about developmental and standard course structures may be difficult and time consuming. Extensive work was required to construct files suitable for analyses from the files provided by the two institutions. The data were organized to suit the institutions' record-keeping needs, which were different for each institution. The data set for Institution I included less detailed information than that for Institution II, and was easier to manage. On the other hand, term-by-term course information provided by Institution II allowed more orderly selection of developmental and non-developmental samples. However, this data set required even more extensive selection, editing, concatenation, and matching to prepare it for analysis. Given these constraints, each application of this methodology would require "tailoring" it to the institution's unique course structure and record-keeping format.

Implications for Postsecondary Institutions

The results of this study have important implications for both educational policy makers and postsecondary institutions that provide developmental instruction. As postsecondary institutions are challenged by policy makers to justify developmental instruction, the approach suggested in this paper can help them document the effectiveness of individual developmental courses. This methodology can also be used to compare the benefits of competing developmental courses and to evaluate tiered developmental course structures. If the purpose of developmental

instruction is to increase the likelihood of future academic success of underprepared students, then courses that seem to provide little or no benefit to these students should be reassessed. These courses should be examined for content relevance and sequencing with subsequent courses.

This approach can also be useful for evaluation of the effectiveness of the institution's overall developmental instruction program in improving student retention and degree completion.

The apparent importance of developmental course grades suggests that postsecondary institutions need to examine their policies for placement into standard courses. This study showed that the commonly-held assumption that earning a C in a developmental course means that a student is prepared for a standard course is not necessarily true. If the results of this study are replicated at other, different types of institutions, they would suggest that institutions should make earning high grades in developmental courses a prerequisite for enrolling in subsequent courses. The methodology laid out in this paper may help determine what the required grades should be.

References

ACT (2003). *ACT Course Placement Service*. Author.

Chen, S. & Cheng, D. X. (1999). *Remedial education and grading: A case study approach to two critical issues in American higher education*. A research report submitted to the Research Foundation of the City University of New York.

McCabe, R. H. (2000). *No one to waste.* Washington, DC: Community Press.

National Center for Education Statistics (1996). *Remedial education at higher education institutions in fall 1995* (NCES 96-584). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

National Center for Education Statistics (2000). *The condition of Education 2000* (NCES 2000-062). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

National Center for Education Statistics (2003). *Digest of education statistics 2002* (NCES 2003-160). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Phipps, R. (1998). *College remediation: What it is, what costs, what's at stake*. Washington, DC: Institute for Higher Education Policy.

Sawyer, R. (1997*). Indicators of the effectiveness of developmental course placement systems in college*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago.

Sawyer, R., & Schiel, J. (2000). *Posttesting students to assess the effectiveness of remedial instruction in college*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans, LA.

Schoenecker, C., Bollman, L., & Evens, J. (1996). *Developmental education outcomes at Minnesota Community Colleges*. Paper presented at the Annual Forum of Association for Institutional Research, Albuquerque, NM.

Spann, M. G. Jr. (2000). *Remediation: A must for the 21st-century learning society.* Policy paper. Denver, CO: Center for Community College Policy.

Weismann, J., Silk, E., & Bulakowski, C. (1995). *Assessing developmental education through student tracking*. Paper presented at the Annual Forum of Association for Institutional Research, Boston, MA.