An Initial Analysis of Student Outcomes Based on DFWs in English 1010, English 1011, and Math 1130

August 14, 2019

This exploratory analysis was produced by OPEIR in response to a request by UTC Academic Affairs. Results are not necessarily generalizable and attempts to use results outside the scope of this project should be avoided.

Introduction and Population Definition

The purpose of this study is to investigate whether UTC students' performance in General Education Gateway courses, such as English Composition and College Algebra, has an effect on long-term student outcomes, taking into account other potential factors. For this initial analysis, the courses of interest are English 1010, English 1011, and Math 1130. The primary outcomes measured are one-year retention rates and six-year graduation rates.

Since one-year retention and six-year graduation are cohort-based metrics, the population of interest is first-time freshmen, and data availability and reliability allow us to begin with the Fall 2011 incoming freshman cohort. For retention this gives us seven cohorts (2011–2017) to work with, and for graduation we have three (2011-2013). A breakdown of these cohorts is in Table 1. The presence of valid high school GPA and ACT score data is important for analyzing freshman cohort performance so, from this point forward, we will refer to the 14,387 students with such data as the *study population*.

Table 1: Incoming Freshman Cohorts and the Study Population

Cohort	All Enrolled Students	Study Population: Students with both GPA and ACT data	Students with grades in English 1010/1011	Students with grades in Math 1130
2011	2,185	2,064	1,397	757
2012	2,284	2,193	1,477	607
2013	2,335	2,251	1,441	1,028
2014	2,144	2,061	1,181	848
2015	1,852	1,786	1,029	785
2016	2,066	1,995	1,076	783
2017	2,134	2,037	1,130	853
Total	**15,000**	**14,387**	**8,731**	**5,661**

Student course performance is treated as binary, with students flagged as either "pass" (grades A, B, C) or "DFW" (grades D, F, W). Grades of incomplete are not considered, and students who registered for but dropped a course without a W are excluded as well. Additionally, we will be looking only at the student's grade in his or her first attempt in the course, and only if that first attempt occurs within the first academic year. DFW rates for these freshman cohorts in their initial attempt in the courses of interest are shown in Table 2.

Table 2: DFW Rates among Study Population for First Attempt in Courses

Cohort	English 1010	English 1011	Math 1130
2011	17%	*	25%
2012	16%	12%	22%
2013	19%	19%	36%
2014	15%	17%	25%
2015	16%	21%	23%
2016	19%	19%	29%
2017	16%	14%	25%
Overall	**17%**	**17%**	**27%**

** English 1011 was first offered in Fall 2012*

Retention and graduation rates for the study population are shown in Table 3, below. Note that the definitions for these outcomes are the same as those used by OPEIR for Factbook reporting (https://www.utc.edu/planning-evaluation-institutional-research/factbook/index.php); one-year retention indicates that the student is enrolled in the following (second) Fall term, and six-year graduation indicates that the student has been granted his or her degree before the seventh Fall term. Since the study population is a very large subset of the overall incoming cohort, it should not be surprising that the rates below match those reported in the Factbook perfectly.

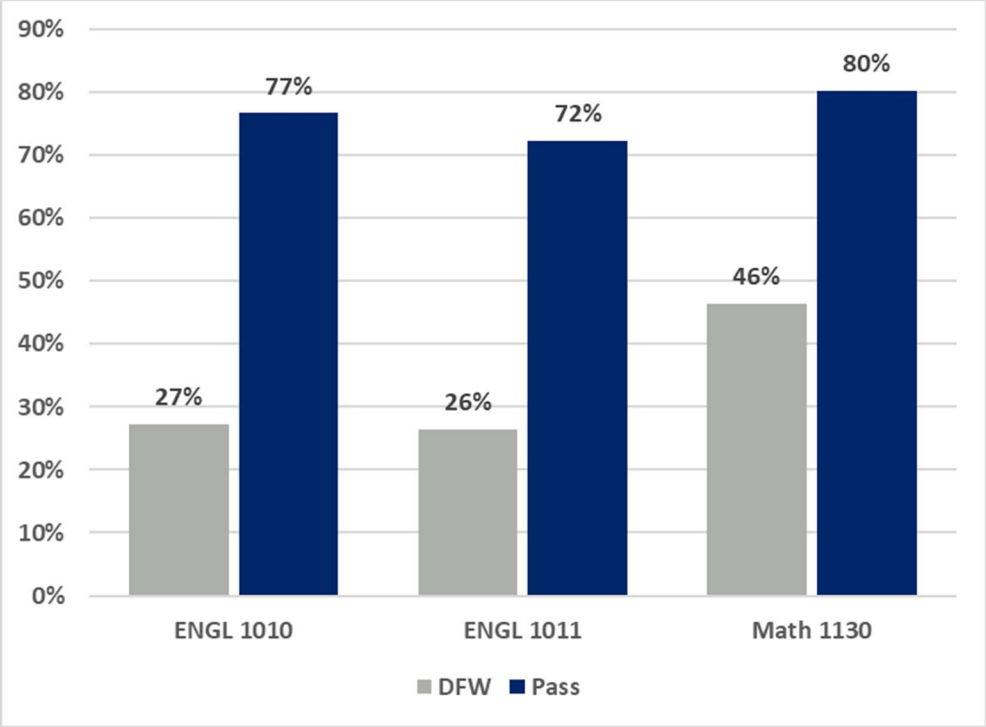
Table 3: Overall Retention and Graduation Rates for Study Population

Cohort	1-year retention	6-year graduation
2011	67%	45%
2012	69%	48%
2013	70%	48%
2014	71%	
2015	74%	
2016	73%	
2017	73%	

Bivariate Analysis of Student Outcomes

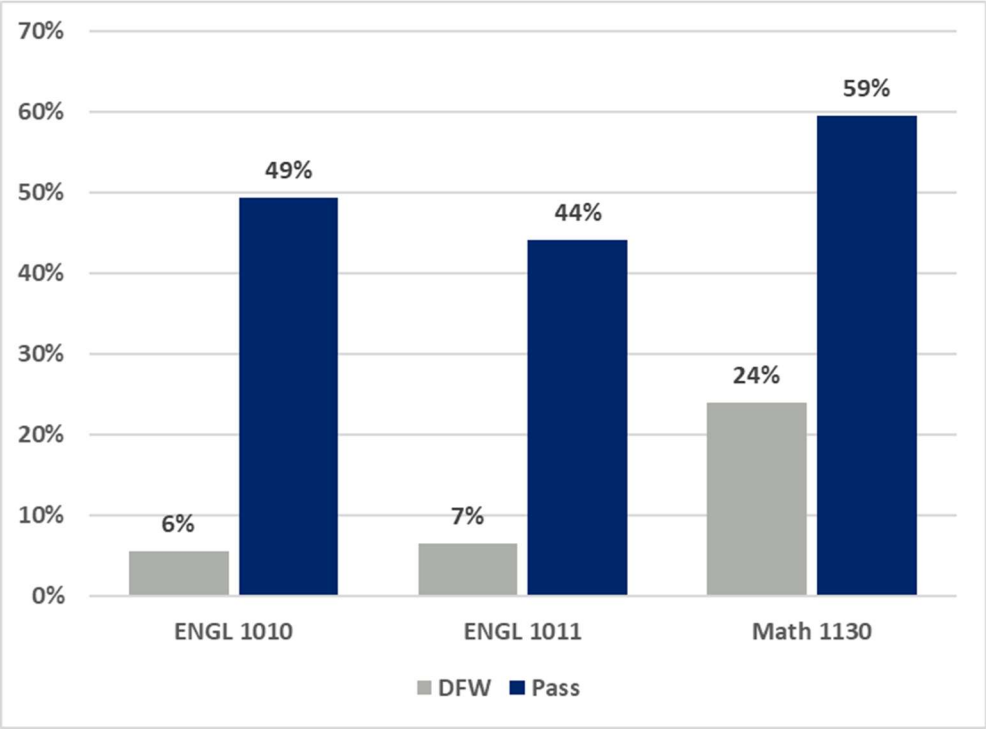
The simplest approach to the relationship between student performance in the courses of interest and the longer-term student outcomes is to break the population into *DFW* and *Pass* groups and observe the student outcomes over time using a series of crosstabs. For each of the three courses, it is clear that students who pass are far more likely to be retained to the following Fall and to graduate within six years. Less than 50 percent of students who received a grade of D, F, or W in any of the courses have been retained in the last seven cohorts, and graduation rates were under 25 percent in the 2011-2013 cohorts. Figures 1 and 2 show the disparity in outcomes by student grade for the three courses.

Figure 1: One-Year Retention Rates for Combined 2011-2017 Cohorts*, by Course Grade



** 2012-2017 for English 1011, since the course was first offered in Fall 2012*

Figure 2: Six-Year Graduation Rates for Combined 2011-2013 Cohorts, by Course Grade*



* *2012-2013 for English 1011, since the course was first offered in Fall 2012*

From this simple analysis two things are apparent. First, students who do not earn a C or better in any of these introductory courses are far less likely to succeed at UTC than those who do. Second, it appears that this effect is far stronger in the English Composition courses than it is for College Algebra. It is worth keeping in mind, however, that DFW rates are significantly higher for College Algebra, so the reduced difference in outcomes still represents a substantial difference in the number of students retained and graduated.

Controlling for Confounding Factors

The preceding analysis confirms that students who do not pass their first attempt at these courses are also much less likely to have good longer term outcomes at UTC. But almost certainly both the student's performance in that initial course and his or her longer term outcomes are related to underlying characteristics of the student or the student's experience. Controlling for these potential confounding factors can give us a better idea of whether performance these early courses, in and of itself, has an independent effect on student outcomes.

The following variables all show a bivariate association with retention and graduation, and were included in a preliminary analysis:

- High School GPA
- ACT composite
- ACT English subscore
- ACT Math subscore
- First generation college student (Y/N flag)
- Adult learner (Y/N flag)
- Race
- Gender
- Family Income (150% poverty level Y/N flag)
- Initial UTC major, department, and college

Many of these variables have strong associations with each other as well, so the relationship of some of them to student outcomes is likely to be spurious. To help determine which among them are truly factors affecting student outcomes and which are not, we will use two different avenues of modeling and variable selection: logistic regression and classification trees.

Logistic Regression with Stepwise Variable Selection

With the binary success/fail outcomes we are analyzing, logistical regression is a sensible analytical approach that will, at least, allow us to see which variables are significant predictors and which are merely correlates of those. With the number of immeasurable factors that must be involved in student retention and graduation we would not expect a model with a tight fit or terrific predictive accuracy, but we can get a good idea of the relationship and predictive nature of the independent variables. With the stronger relationship between English Composition DFWs and our outcomes, we will, for now, focus on students who took English 1010 or 1011 in their first year. Many of them also took Math 1130, so we can still include performance in that course as a possible factor.

Modeling One-Year Retention

Stepwise selection results in a model that has seven variables that register as statistically significant, but it appears that some are the result of overfitting, as they do not improve classification accuracy at all and some in fact have counterintuitive signs. Table 4 shows the results of variable selection.

Table 4: Stepwise Model Selection Results for One-Year Retention

Selection Step	Variable added	Wald Statistic in Full Model	Model Chi-Squared	Classification Accuracy
Step 0	(Constant)	21.9	-	67.2%
Step 1	English Comp DFW	724.1	1233.0	75.0%
Step 2	College Algebra DFW	91.0	1373.4	75.0%
Step 3	High School GPA	75.6	1449.8	74.9%
Step 4	First Gen	39.0	1488.7	74.9%
Step 5	Race	29.1	1517.9	75.0%
Step 6	ACT Math	9.3	1528.1	75.1%
Step 7	Department	42.8	1572.6	74.0%

In the final model all seven variables are statistically significant, but after *English Comp DFW* has been added no additional variables improve classification accuracy. We also see only marginal improvements in the model chi-squared statistic, and every variable has a significantly smaller Wald statistic (essentially a partial Chi-square for each variable in the model) than *English Comp DFW*. Whatever explanatory power these variables have alone, none of them add any explanatory power to a model that already includes *English Comp DFW*.

We should consider that the effects of several of these variables, if they truly are predictive, are already at play in the student earning the DFW. A student who is academically less ready (low high school GPA or ACT score) or disadvantaged in some other way (First Gen, Race, Family Income) may be far more likely to DFW early courses. Nonetheless, we see tremendous predictive power in *English Comp DFW*, and no additional information is needed to provide the best success prediction possible in this model.

Modeling Six-Year Graduation

Stepwise selection again results in a model that has seven variables that register as statistically significant, and again the bulk of the improvement in classification accuracy comes from the initial step adding *English Comp DFW.* Most of the variables selected are the same as in the retention model, including the first four. Table 5 shows the results of variable selection for the graduation model.

Table 5: Stepwise Model Selection Results for Six-Year Graduation

Selection Step	Variable added	Wald Statistic in Full Model	Model Chi-Squared	Classification Accuracy
Step 0	(Constant)	116.3	-	40.9%
Step 1	English Comp DFW	204.6	549.6	55.9%
Step 2	High School GPA	78.5	688.4	57.3%
Step 3	College Algebra DFW	40.0	733.9	58.6%
Step 4	First Gen	24.8	763.9	59.6%
Step 5	ACT Math	20.8	778.8	59.6%
Step 6	College	14.3	799.7	60.2%
Step 7	Major	64.4	873.6	61.5%

In the final model all seven variables are again statistically significant. This time we do see some improvement in classification accuracy for variables added after *English Comp DFW* but we see only marginal improvement in the model chi-squared statistic after Step 3, and again every variable has a significantly smaller Wald statistic than *English Comp DFW*. One speculative, but logical explanation is that for six-year graduation, there is much more time for other factors to play a role and reveal some significance.

What is clear from these models is that *English Comp DFW* is very strongly associated with both one-year retention and six-year graduation, even when accounting for other variables that we would expect to be significant. It is also interesting to note that *College Algebra DFW* is among top three variables in both models as well, suggesting that performance in this course may have a relationship that is less powerful but nevertheless somewhat distinct from performance in English composition. The other variables which persist, at least in terms of statistical significance, are two measures of academic readiness (*High School GPA* and *ACT Math*) and one representing non-academic disadvantage (*First Gen*). After further analysis, we will see that *High School GPA* may be the most important factor that is known before the student begins coursework.

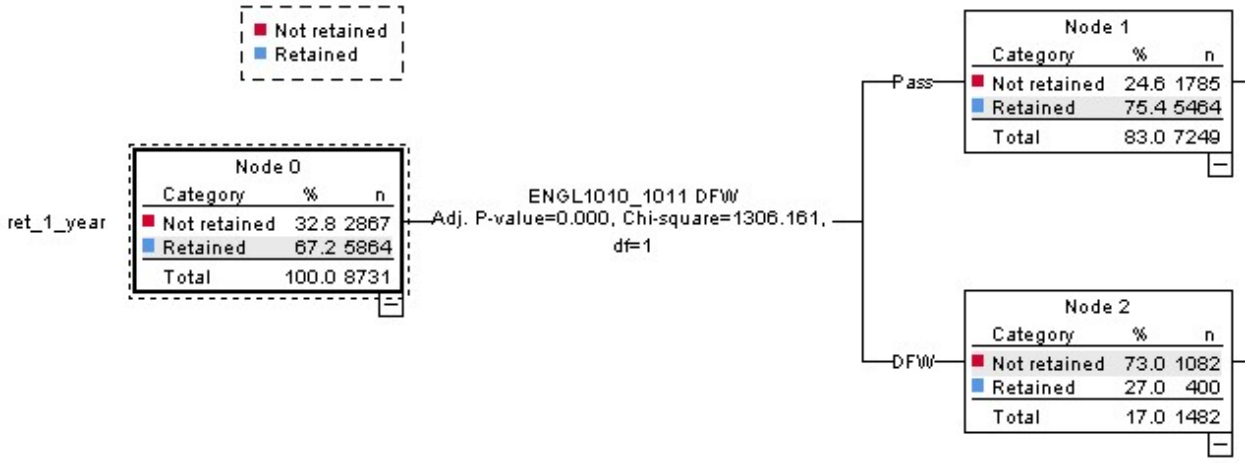
Classification Trees

Another analytical approach to determine the association between a number of variables and an outcome of interest is the use of a Classification (or Decision) Tree. This has the advantage of being much easier to interpret than a logistic regression model and shows in the form of a visual hierarchy of importance the various predictors that may be statistically significant. Through the use of SPSS and its CHAID (Chi-Square Automatic Interaction Detection) routine, we can look again at the impact of English and Math DFW grades in the presence of other possibly predictive variables. For ease of interpretation, several quantitative variables (High School GPA and ACT Composite and subscores) have been binned into quartiles. This is not necessary for CHAID to work but it improves interpretability of results.

Retention Classification Tree

As shown in Figure 3 below, the initial split from the initial population occurs using the grade earned in English Composition. Among the 8,731 students in Node 0 (the available population from cohorts 2011-2017), just over 67% were retained. But when you split this group into those who pass ENGL 1010 or 1011 and those who DFW the course, you get two groups that look dramatically different: over 75 percent retention for those who passed compared to just 27 percent for those who did not.

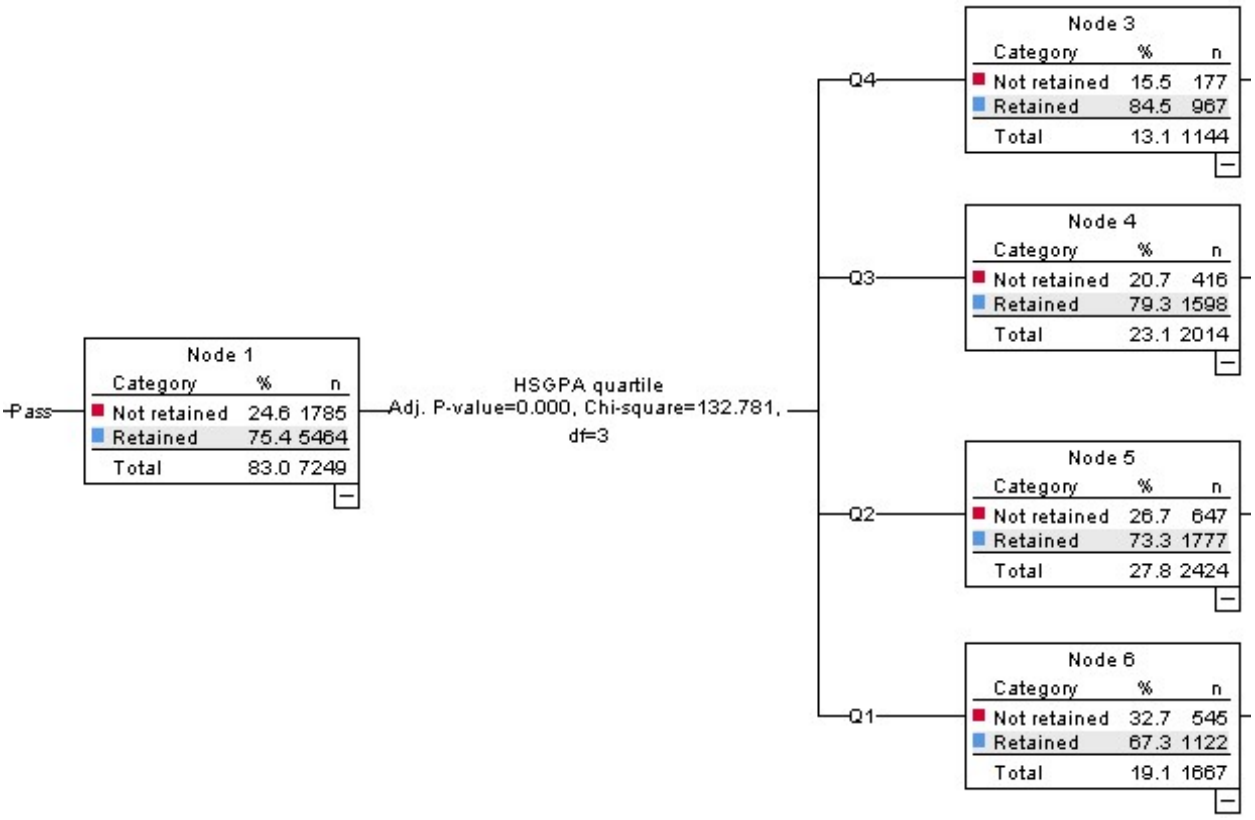
Figure 3: Initial Split in Retention Tree



As with the logistic regression model on the same data set, we find that after English Composition grade is taken into account, all other variables show significantly less predictive power. Each side of the tree does split for two more levels but the largest Chi-Square statistic is an order of magnitude smaller than that associated with the first split.

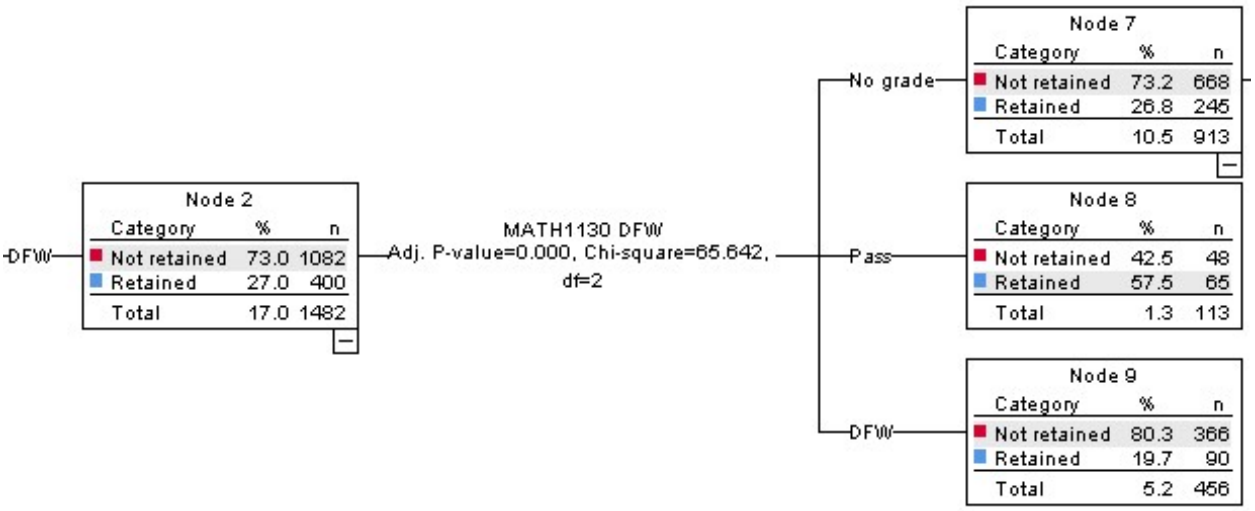
Figure 4 shows the next split on the upper portion of the tree, where students who passed English Composition are split by the next most significant variable, High School GPA. Not surprisingly, the four quartiles of GPA are associated positively with retention, with the bottom quartile retained at just a 67 percent rate and the top retained at nearly an 85 percent rate. This is an intuitive result but it is important to note that among this group, even the worst-performing quartile is retained at a rate more than double those who did not pass English Composition, and the Chi-Square statistic associated with this split is considerably smaller.

Figure 4: Second Split in Retention Tree with Passing Grade in English Comp



On the bottom of the tree, we observe a different factor making an appearance. As seen in Figure 5, among students who did not initially pass English Composition the next most predictive variable is the student's performance in College Algebra.

Figure 5: Second Split in Retention Tree with DFW in English Comp



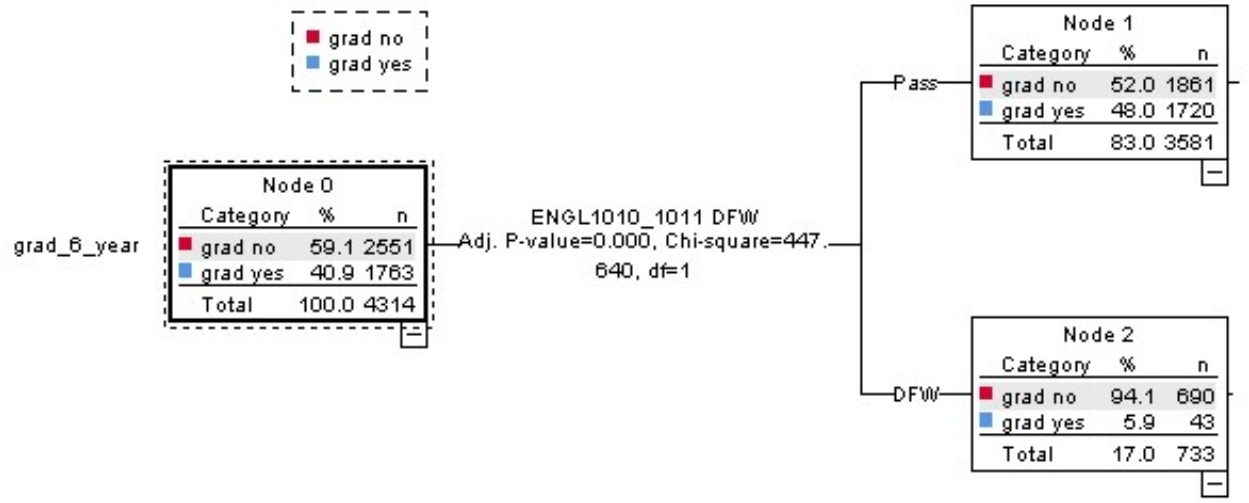
While it is interesting to observe that there has not been much difference in retention between those who earned a DFW in College Algebra and those who did not take it, the more interesting result is in Node 8. While there are not many students who earned a DFW in English Composition but passed College Algebra in their first attempt, those who fit into this group are twice as likely to be retained as those who did not pass (or take) College Algebra. This suggests that perhaps there is something noteworthy about passing *either* course that makes a student more likely to return to UTC for their second year.

Nodes 3 through 7 all split one additional time, but at this level both effect significance and interpretability are greatly diminished. It is significant to note, however, that for each GPA quartile among those who passed English (Nodes 3 through 6), the next split is based on Math performance; for the bottom three quartiles the split is on College Algebra DFW, and for the top quartile it is on ACT Math subscore (largely because fewer of these students took College Algebra at all). In all cases, earning a DFW in College Algebra was associated with lower retention rates, though even in the worst case retention was above 65 percent.

Graduation Classification Tree

As we may have expected based on the analysis so far, English Composition DFW is the most predictive factor for six-year graduation as well. Figure 2 showed that students who did not pass their first attempt at English 1010 or 1011 have graduated at only a six percent rate—just 43 graduates among 733 students. Figure 6 shows that this is indeed the most predictive factor for long term success among all of the variables analyzed.

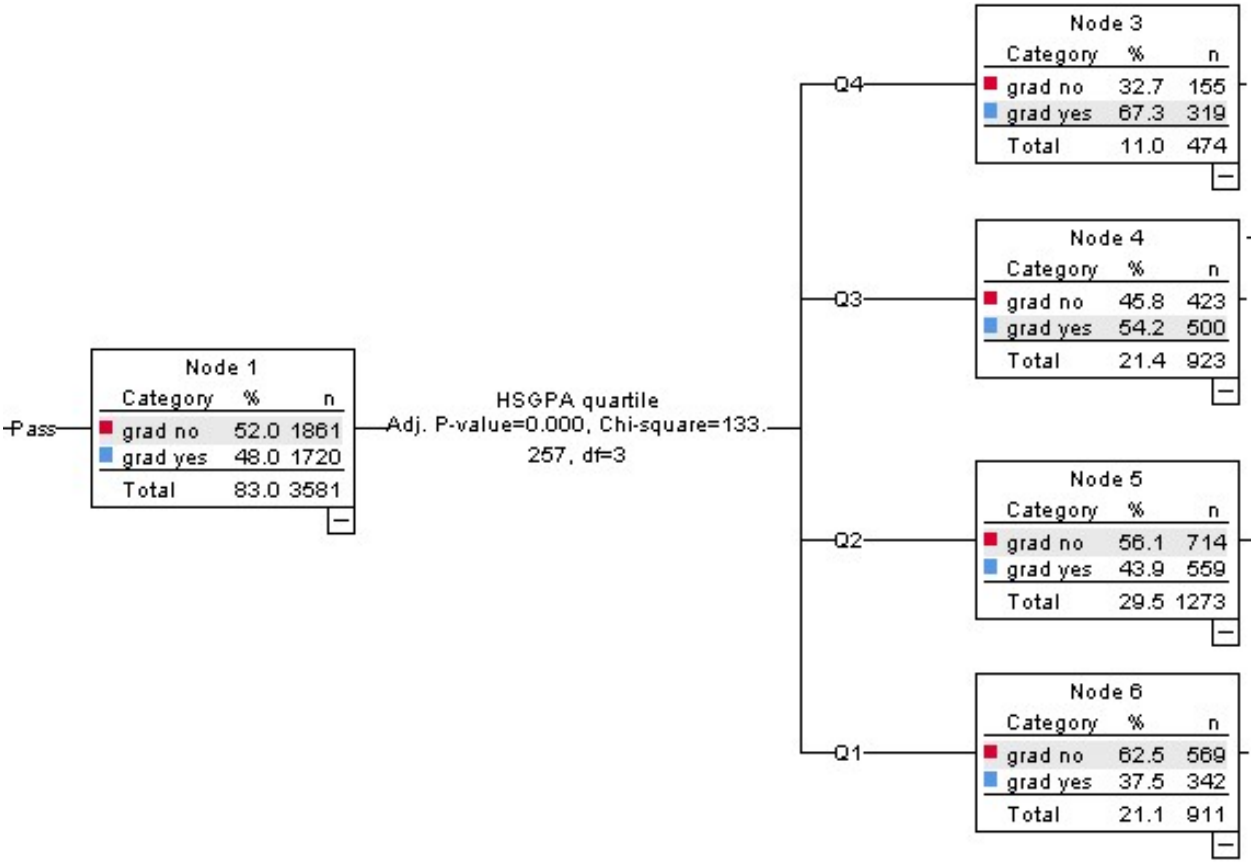
Figure 6: Initial Split in Graduation Tree



In the next split, the top of the tree brings in High School GPA, again not surprising based on the retention results and on the logistic regression model. Unlike the retention tree, however, it

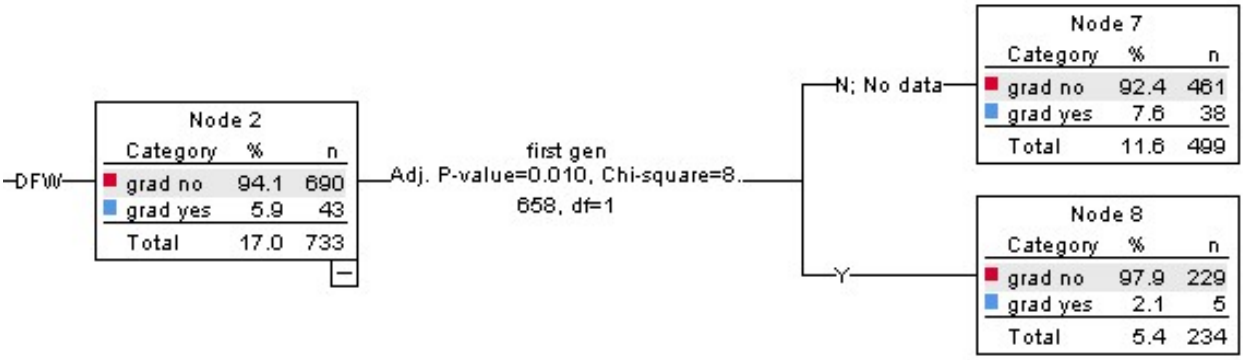
plays a stronger role in classifying students by likelihood of graduation (Figure 7). Among students who passed English Composition, the graduation rate is nearly twice as high for those in the top quartile of High School GPA as it is for those in the bottom quartile. This again suggests that High School GPA might be the best predictor of long term success *before* a student sets foot in a UTC English or Math classroom.

Figure 7: Second Split in Graduation Tree with Passing Grade in English Comp



In the bottom of the graduation tree we only see one additional split, and that is based on the First Generation College Student flag (Figure 8). While the difference in percentage points and the Chi-Square statistic are small compared to other factors, it is still statistically significant and potentially meaningful for longer term student success. Small numbers aside, it is interesting to note that students self-identifying as First Gen on the FAFSA have been more than three times less likely to graduate in six years than those who are not First Gen if they do not pass their first attempt at English Composition. It is pretty remarkable that among 234 such students in the 2011-2013 cohorts, only five graduated.

Figure 8: Second Split in Graduation Tree with DFW in English Comp



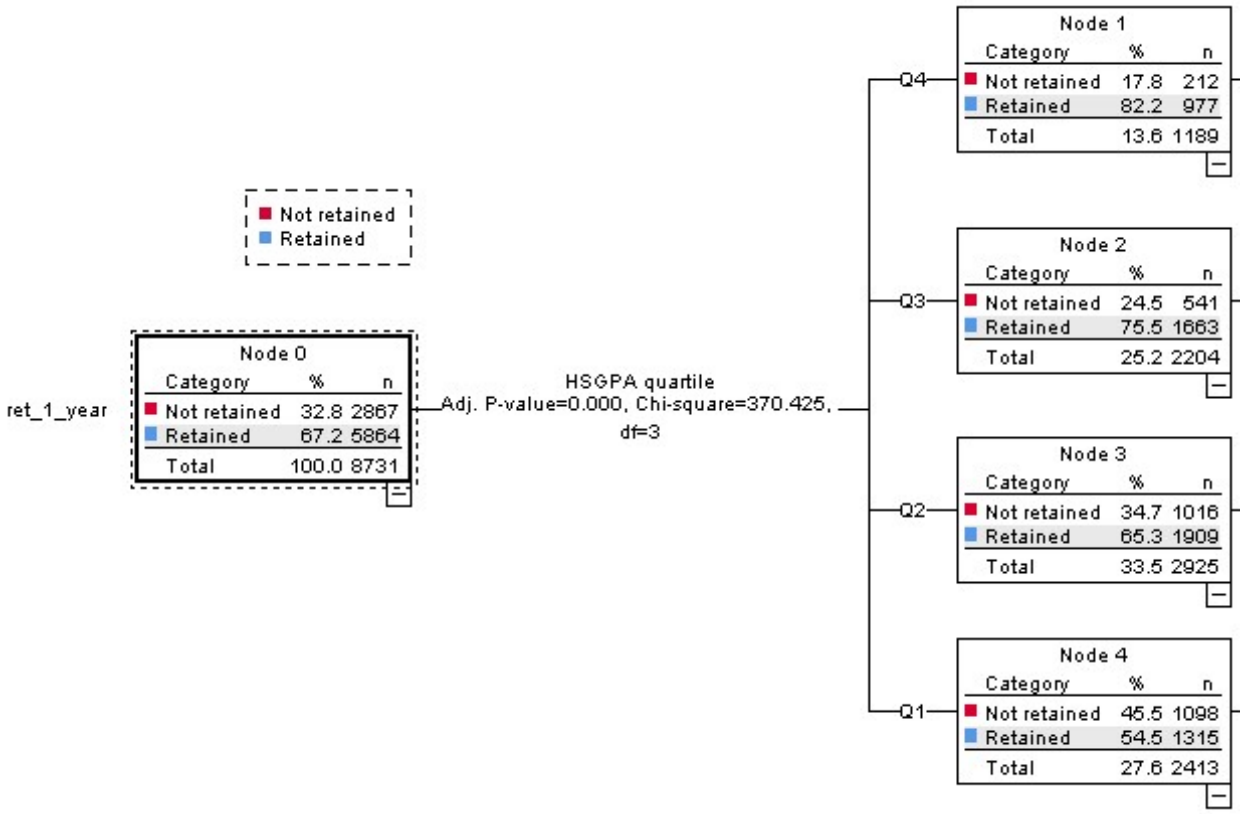
In the top of the tree, Nodes 3 through 6 split an additional time, but the results are inconsistent and statistical significance diminished. Among students in the lowest quartile for High School GPA, College Algebra grade shows up as the most significant remaining variable, with a passing grade in College Algebra improving the chance of graduation. Meanwhile, the top three quartiles show non-academic disadvantage coming into play—low family income for the second quartile and First Gen status for the third and fourth quartiles. In each case, the disadvantage is associated with a reduction in graduation rate.

The Role of High School GPA: Explicit Control in Retention and Graduation Trees

In both trees it is clear that success in English Composition is a strong predictor of longer term student success, but how much of this is just a reflection of overall college-readiness of students? Unlike in regression models, classification trees don't apply controls explicitly, and don't use variables as controls at all unless they control a split at the beginning of the tree. Since both intuitively and analytically it is clear that High School GPA plays a role in both one-year retention and six-year graduation, it is important to make sure that English 1010 and 1011 grades are still significant predictors after explicitly controlling for it. To do this, we force High School GPA to be the first variable in the tree to see what, if any, variables split the tree again.

Figure 9 shows the first (control) split of the retention tree, which shows that High School GPA alone is a very strong predictor of one-year retention. Students in the top quartile have been retained at a rate of 82 percent, while those in the bottom quartile have been retained at just a 55 percent rate.

Figure 9: High School GPA Control Split in Retention Tree



Significantly, the next splits in the tree show that performance in English Composition remains an important predictor of success even after controlling for High School GPA (Figure 10). In all but the top quartile of High School GPA the next split is on English 1010 or 1011 DFW, showing that grade in that course is important even after adjusting for academic readiness. For each the first through third quartiles we see a retention rate for students passing English more than double that of those earning grades of D, F, or W. In the top quartile, we instead see First Gen as the most significant variable, though the difference in retention rates is not nearly as dramatic and the Chi-Square statistic is much smaller.

Figure 10: First Split after Controlling for High School GPA in Retention Tree

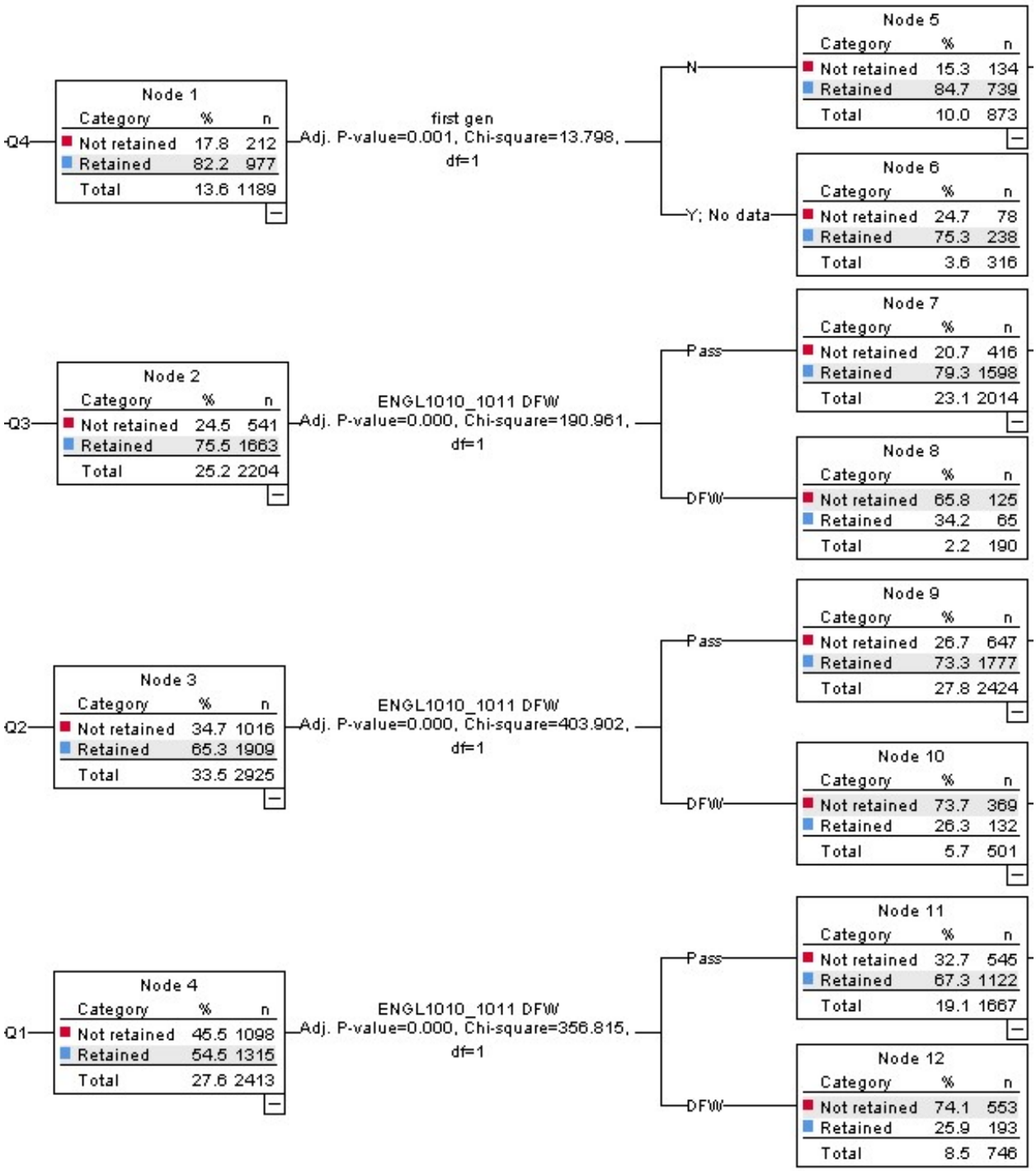
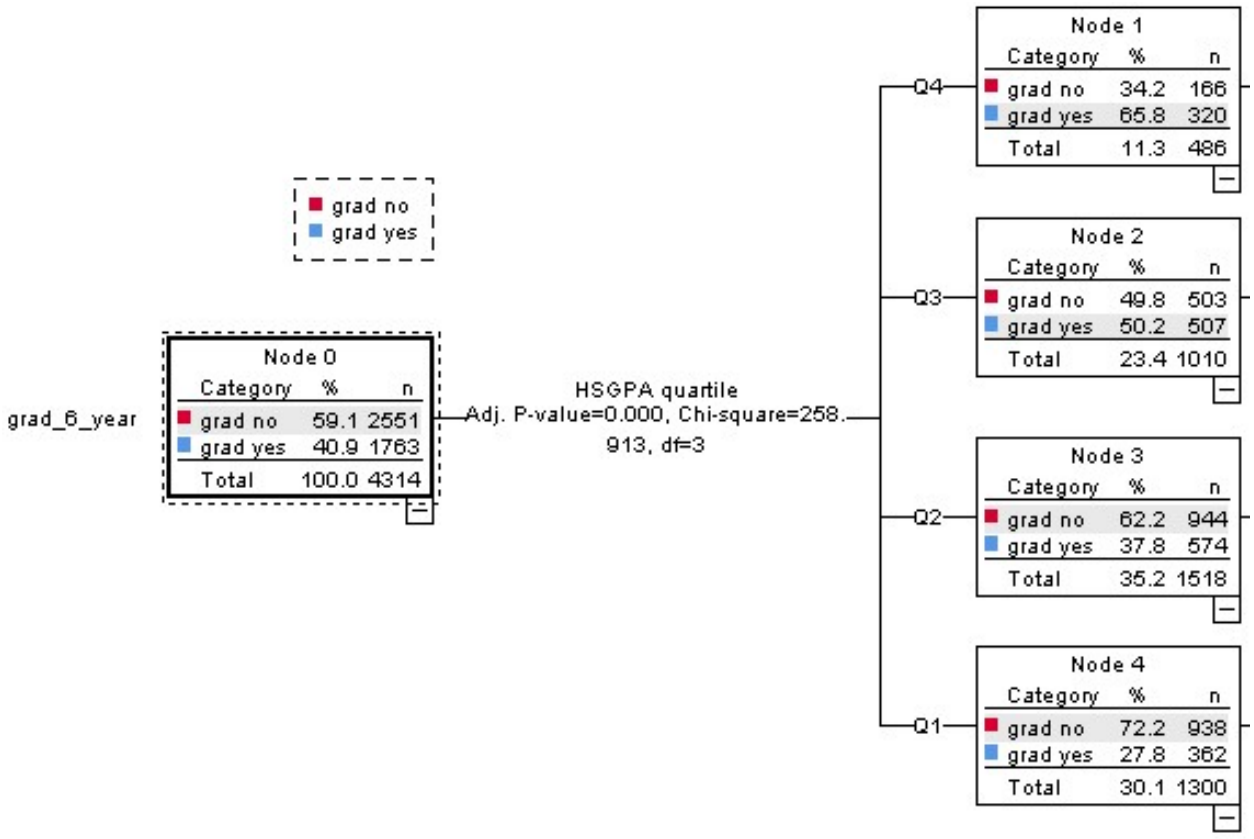


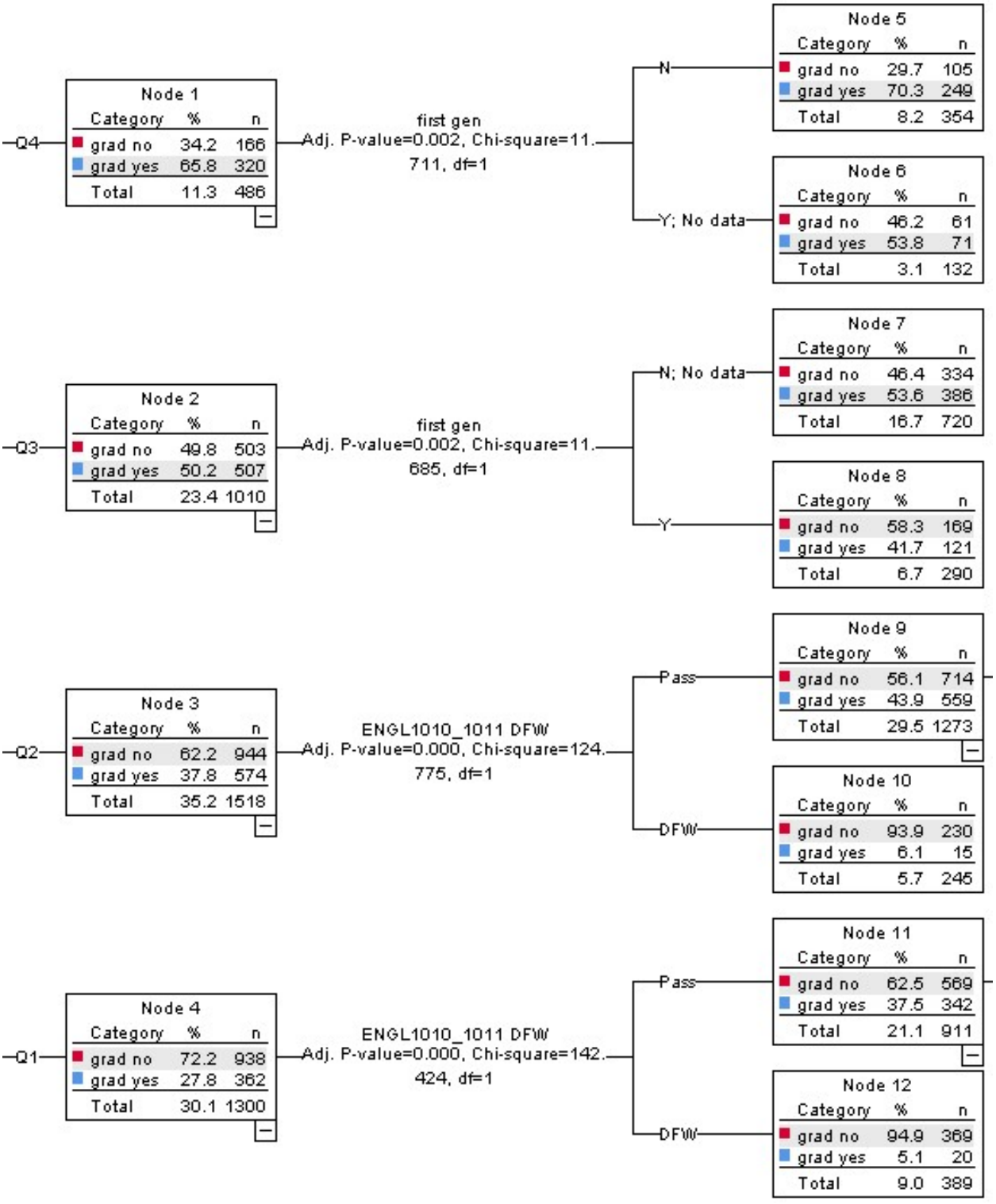
Figure 11 shows the first (control) split of the graduation tree, where again we see a strong association between High School GPA and long-term student success. Students in the top quartile have graduated at rates more than double those of students in the bottom quartile.

Figure 11: High School GPA Control Split in Graduation Tree



As in the retention tree, we see that English Composition grade and First Gen status are the most significant predictors of success after controlling for academic readiness (Figure 12). This time, the effect of being a First Generation student is the most significant for the top two quartiles, while passing English 1010 or 1011 is the best predictor for the bottom two quartiles. Once again the effect of First Gen is much more subtle, with a Chi-Square statistic of about 11 for each of the top two quartiles, while the effect of English grade is very substantial, with Chi-Square statistics more than ten times as large. In each of the bottom two quartiles, we see that students who pass English Composition are about seven times more likely to graduate within six years than those who do not. The magnitude of this effect is quite remarkable considering that High School GPA has already been used to adjust for overall academic readiness.

Figure 12:


Node 1
Category % n
grad no 34.2 166
grad yes 65.8 320
Total 11.3 486
Q4
first gen
Adj. P-value=0.002, Chi-square=11.711, df=1
N
Y; No data
Node 5
Category % n
grad no 29.7 105
grad yes 70.3 249
Total 8.2 354
Node 6
Category % n
grad no 46.2 61
grad yes 53.8 71
Total 3.1 132
Node 2
Category % n
grad no 49.8 503
grad yes 50.2 507
Total 23.4 1010
Q3
first gen
Adj. P-value=0.002, Chi-square=11.685, df=1
N; No data
Y
Node 7
Category % n
grad no 46.4 334
grad yes 53.6 386
Total 16.7 720
Node 8
Category % n
grad no 58.3 169
grad yes 41.7 121
Total 6.7 290
Node 3
Category % n
grad no 62.2 944
grad yes 37.8 574
Total 35.2 1518
Q2
ENGL1010_1011 DFW
Adj. P-value=0.000, Chi-square=124.775, df=1
Pass
DFW
Node 9
Category % n
grad no 56.1 714
grad yes 43.9 559
Total 29.5 1273
Node 10
Category % n
grad no 93.9 230
grad yes 6.1 15
Total 5.7 245
Node 4
Category % n
grad no 72.2 938
grad yes 27.8 362
Total 30.1 1300
Q1
ENGL1010_1011 DFW
Adj. P-value=0.000, Chi-square=142.424, df=1
Pass
DFW
Node 11
Category % n
grad no 62.5 569
grad yes 37.5 342
Total 21.1 911
Node 12
Category % n
grad no 94.9 369
grad yes 5.1 20
Total 9.0 389

Conclusions

While a large sample size and attempts to control for confounding variables might help tease out correlation from cause, nothing in the analysis here confirms that failing a first English or Math course at UTC *in and of itself* causes worse long-term student outcomes. However, the analysis does make it quite clear that performance in English Composition, and to a lesser degree performance in College Algebra, is an incredibly useful *indicator* of the likelihood of success for a student. This is the case even after accounting for other useful predictors of student success such as measures of academic readiness (High School GPA, ACT scores) and non-academic factors (family income, First Gen status, race, etc.).

It appears that this is especially true for students with below-UTC-median High School GPA. When explicitly accounting for this measure of academic readiness, student outcomes showed substantial divergence based on English course grade. While we already know student success is tied to readiness, we can now clearly see that early performance in key Gateway courses is especially crucial for those who appear less ready based on academic credentials. Graduation rates under 10 percent for such students serve as a powerful suggestion that such courses, and student performance therein, deserve a very close eye from anyone interested in improving student outcomes and increasing retention and graduation rates. This could be key in moving toward the UT System's ambitious goals in these metrics.

Lingering Questions and Avenues for Additional Research

Closely related topics considered during this analysis and as potential extensions of this work include:

- **Performance in English 1020 (Rhetoric and Composition II) when that is a student's first English course.** Preliminary analysis suggested that retention and graduation is higher among these students, but that earning a D, F, or W in this course may still have predictive power in long term student outcomes.
- **Placement policies for English 1010, 1011, and 1020.** Directed self-placement is used for this, and students are currently guided to the appropriate course based on ACT English subscore. Initial analysis has suggested that this score actually has very little predictive power for success in English Comp courses at UTC, regardless of where a student is placed, and that High School GPA may be a better guide. It was evident that the extra assistance for students enrolling in English 1011 is effective in improving outcomes, but the question remains whether we are guiding the *right* students to this course.
- **Impact of performance in other Math courses.** The analysis of Math grades was weakened by the fact that the most academically ready students are far more likely to be taking more advanced math courses for their General Education and/or major requirements. Similarly, students who elect to take Math 1010 (Mathematics in Our

Modern World) instead of College Algebra may show significantly different long-term outcomes.

- **Other measures of English and Math readiness.** Preliminary work showed that High School GPA was the academic readiness metric that had the most predictive power for student success, and even combined with the various ACT scores and subscores very little additional power was achieved. However, results of other studies suggest that *relative* strength in certain subjects may be as predictive as absolute strength due to self-efficacy and other effects. For example, it may be worth investigating whether the ratio of ACT English subscore to Math subscore or Composite score has better predictive power than ACT English alone.
- **Projecting the likelihood of DFW in English or Math.** If there is a causal relationship between an English or Math DFW itself and worse long-term outcomes, then this would suggest that certain at-risk students receiving extra help in English Comp or College Algebra could improve their chances of success. It may be useful to establish a risk indicator specific to performance in those courses.
- **Fall-to-Spring retention of students earning a DFW in English Comp or College Algebra.** If students aren't still enrolled in their second term, then intervention will not even be possible, much less effective, in trying to keep them enrolled through their second Fall. It is important to see how quickly we are losing these students.
- **Instructor and scheduling influence on grades in these courses.** If we are seeing a causal relationship between performance in these Gateway courses and long-term student outcomes, then assigning quality instructors and scheduling course meetings at times which maximize success would take on even greater importance.
- **Effective interventions in second term or mid-first term.** This is outside the scope of this line of analysis, but the next logical step would be investigating what types of actions could be taken by UTC instructors, advisors, departments, or student success staff to mitigate the effects observed here—to prevent grades of DFW, if possible, or to assist academically those earning DFWs to keep them engaged at UTC.