

# Analysis of DFW Rates for the Fall 2013-Spring 2017 CSU Chico Classes

Jeff Bell

Institutional Research Faculty Fellow

Poor grades, especially in lower division courses (LD), are responsible for the Underrepresented Minority (URM) and gender gaps in graduation rates (see “Analysis of Dismissal and Dropout Rates”) and are also responsible for a significant fraction of the students lost to graduation. This analysis looks at general factors (student characteristics and general class characteristics) that affect whether students receive poor grades in a class, and how these factors affect the URM and gender gaps. Class and Instructor effects are also important but will be examined in a subsequent report.

## Summary of Results:

- The distribution of student grades is not normally distributed, with the most common grade being an A, ~24% of all grades, and then a steady decline in the proportion of grades with the exception of bumps up for whole B, C, D, and F grades.
  - The odd distribution of the different grades indicates that faculty are not grading on a “curve”, so the basis for the cut-offs for the different grades is unknown and likely quite variable.
  - The proportion of poor grades, DFW (D+, D, F, and WU), was used as a measure of the proportion of students doing poorly in a class.
- While the overall DFW rate for all classes is ~10%, the DFW rate for LD classes, 13%, is 1.73-fold worse than the 7.5% rate in UD classes.
  - The high DFW rate in LD classes explains the high dismissal rate in the first two years for FTF; lowering these high DFW rates will be necessary to improve our graduation rates.
  - This is not caused by the addition of transfer students to the UD classes as transfer students do a little worse in UD classes than students who started at Chico as First Time Freshmen (FTF), 7.6% DFW vs 7.1%, but a little better in LD classes, (12.5% vs 13%), thus their presence decreases the gap between LD and UD slightly.
    - This also indicates that students are not doing poorly in LD classes because of a lack of experience, as transfers, and FTF juniors and seniors taking LD classes do almost as poorly in LD classes as new students.
    - Community College preparation and our current screening procedures provide transfer students that are equivalent to our FTF juniors.

- URM students are 1.55-fold more likely to get a DFW grade than non-URMs and males are 1.53-fold more likely to get a DFW grade than females.
  - These are independent, so URM males are 2.5-fold more likely to get a DFW grade than non-URM females while there is little difference between non-URM males and URM females.
  - Reducing the differences in DFW rates between URM students and non-URMs is critical to reducing the gap in graduation rates.
- The HS eligibility index measure of HS preparation is a strong predictor of DFW rates, with every 0.1 increase in the eligibility index reducing the DFW rate by 1.1 percentage points.
  - FTF entering with index scores below 2.8 have DFW rates of ~25%, over 10-fold higher than the 2% rate for FTF with index scores over 4.4.
  - ~30% of the URM gap can be explained by URM students coming in with average eligibility scores 0.15 less than the non-URMs; the reason for the remaining 70% of the difference is unknown.
  - At eligibility scores over 4.0 there is no URM gap.
  - Very little of the gender difference in DFW rates can be explained by HS eligibility scores.
  - Reducing the persistent URM gap in DFW rates after controlling for HS eligibility scores will be required to reduce our URM graduation gap
- Oddly, GE courses have lower DFW rates than non-GE at the LD level, but higher rates at the UD level
  - Writing Intensive and capstone classes have DFW rates one third lower than other GE classes at both the LD and UD level.
  - The URM DFW rates are ~1.5-fold higher for all types of classes.
  - Interestingly, the gender-based grade gap between males and females narrows in GE classes, especially UD GE, where they are only 1.18 times more likely to get a DFW grade, instead of the usual 1.5 times more likely.
- Larger Classes tend to have higher DFW rates with the largest LD lecture classes, >200 enrolled, having twice the DFW rate of the smallest classes, <11 enrolled.
  - While the increase in DFW rates keeps increasing up to the largest size for LD classes, the increase plateaus around an enrollment of 50 for UD classes.
  - UD classes have DFW rates 30-50% lower than LD classes at all class sizes, with the gap being greatest for the largest sizes.
    - Thus, only about one third of the DFW gap between LD and UD can be explained by smaller class sizes in the UD classes. The cause of the remaining two thirds of the gap is unknown.
  - LD courses have more “large classes” with enrollments over 60, while having fewer “small classes” with enrollments less than 40.
  - Large classes have more of a harmful effect on females, especially URM females.
  - Removing LD Math classes from the LD class size analysis reduced the DFW rate for classes with 31 to 40 enrolled by 71%, showing that there are strong subject effects on DFW rates.

- At the LD level the DFW rate varies by a factor of three between classes in the different Colleges, from a low of 6.7% for CME to a high of 23.6% in Business.
  - The URM gap is similar for classes in each of the colleges.
  - Gender differences do vary some, from a low of a 1.09-fold higher risk of males getting a DFW in Business to a high of 1.72-fold higher in CME, so subject effects on the gender gap appear to be larger than their effects on the URM gap.
    - In classes with a higher percentage of males the relative risk for males getting a DFW is lower than in classes with a high proportion of females.
- Within a single college the DFW rate for Departments can vary by more than four-fold. For instance, in the College of Behavioral and Social Sciences (BSS) they vary from a high of 22.6% in Social Sciences to a low of 5.6% in Health and Community Services
  - These differences are puzzling as there is no obvious reason why LD Psychology classes should be almost twice as difficult (13.2% DFW rates) as Sociology classes (7.3% DFW rate).
  - There is much more variation in the relative differences between URMs and non-URMs at the Department level, with the relative risk of a DFW vary from no difference in Social Work to a high of 2.3-fold more in Sociology
    - Departments where more URMs tended to take classes had a lower relative risk of a DFW for URMs.
  - While not as large as for URMs, there is also wide variation in the relative differences between males and females at the department level, from a low of only 1.13-fold higher in Economics to 1.96-fold in Health and Community Services.
    - As with URMs, and as was seen at the College level, males did better relative to females in departments where more males were taking classes.
- The department with the highest LD DFW rate out of all departments is Accounting, with a DFW rate of 32%. However, Math is responsible for the largest number of LD DFW grades, giving out over 600 a semester.
  - Reducing the LD DFW rate and the relative risk for URMs in the departments giving out the most URM DFWs, such as Math, History, Psychology, Biology, Chemistry, Anthropology and Philosophy is the only way we can raise overall grad rates and reduce the URM gap in grad rates enough to make our GI 2025 targets.

### **Methodology**

The data used for this analysis was a combination of data on grades in all classes for the Fall, 2013 through Spring 2017 semesters (data extracted from CRA #659 by Jeff Bell) with initial enrollment information on student characteristics - ethnicity, gender, initial major, etc. (data extracted from CRA #508 by Jeff Bell), schedule information (extracted from CRA #1388 by Jeff Bell) and course GE status (data provided by Bill Loker). The 8 semesters of grade data consist of 886,634 records, of which 652,469 are grades on an ABC basis. Non-enroll sections (associated labs, activities, etc.) are responsible for 153,481 entries with no grade. There are

also ~38,000 Cr/NC grades and small numbers of other special courses with different grading schemes. For the purposes of this analysis only the 652,469 ABC type grades were used.

The primary dependent variable used to measure student success for this analysis, or the lack thereof, is the DFW rate, the percentage of grades that are D+, D, F, or WU, as these grades all result in GPAs below 2.0 and can lead to dismissal and frequently require the student to retake the course, delaying graduation. Note that drops, Incompletes, and Ws are not included in this dataset. When comparing the risk of getting a DFW grade between URM and non-URM, etc., a relative risk was calculated, the ratio of the probability of getting a DFW grade in one population to the probability in the other population.

## Results

The ABC grades are not distributed normally, instead far and away the most common grade in all courses is an A, 24% of all grades, and there is then a liner fall off from there to the lower grades. The exception to this is the strong preference for “whole” grades, as the “+” and “-” versions of grades are significantly underrepresented. While the excess of As and Fs plus WUs is probably caused by the summing of long tails at either end of the distribution of student performance, the excess of Bs, Cs, and Ds suggests that some faculty are not using + and – grades. As student performance is almost certainly normally distributed, most faculty are not grading on a “curve”, so it is unknown what the basis for the grade cut-offs is. Because ABC grades are not normally distributed, average GPA will not be a very useful method for comparing grades between students or classes.

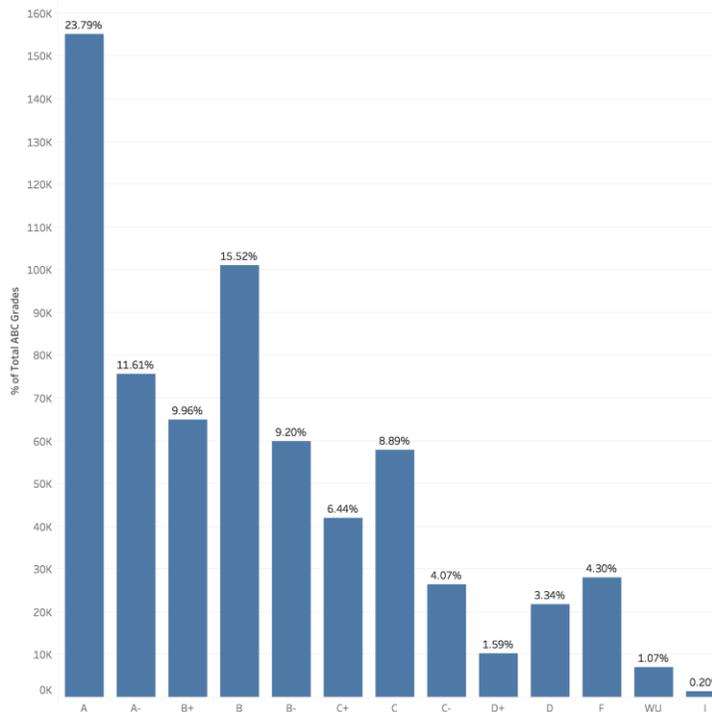


Figure 1. Distribution of ABC grades for the Fall, 2013 to Spring 2017 semesters

For the rest of this analysis DFW rates will be used as the primary measure of student under performance. For all lower division (LD) and upper division (UD) courses, the average DFW rate is ~10% with 67,000 poor grades for the 8 semesters, or 8,375 poor grades a semester. Comparing transfer students to students who started as First Time Freshmen (FTF) in LD and UD courses, Fig. 2, shows that transfers are comparable to FTF, doing a little worse in UD courses (7.6% DFW vs 7.1%), but a little better in LD courses (12.5% vs 13%). To keep sample sizes as large as possible, most of the following analyses will combine FTF and Transfers. As expected from the dismissal data in “Analysis of Dismissal and Dropout Rates”, LD DFW rates, 13%, are much higher than UD rates, 7.5%, so some analyses will use only LD rates as this is the largest problem. The relative risk of getting a DFW in a LD class is 1.73-fold higher than in an UD class.

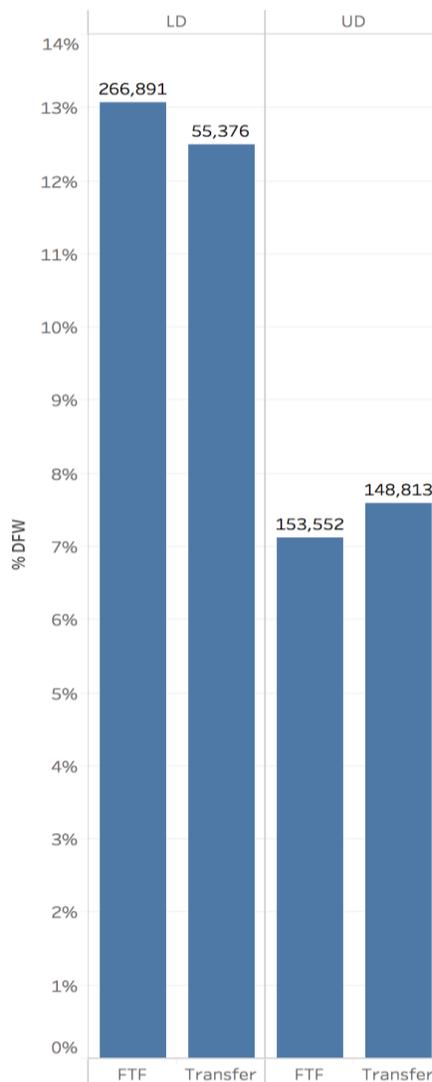


Figure 2. DFW Rates by Class Level and by Admit Type. The Admit Type category includes grades for all students who started as FTF or as Transfers.

As expected from the graduation rate and dismissal rate analyses, there are URM and gender gaps in the DFW rates and these are independent, Fig. 3. The URM DFW rate of 14.5% is 55% higher than the 9.3% rate for non-URMs, while males' DFW rate of 13.5% is 53% higher than females 8.8% rate. Non-URM females have the lowest DFW rate, 6.9%, while URM males have the highest rate, 17.4%, 2.5-fold higher than for white females. Non-URM males and URM females have intermediate DFW rates of 12% and 12.5%, respectively. If the URM and gender effects are independent we would expect the URM male rate to be 1.55 (URM) x 1.53 (male) = 2.37 times higher than the non-URM female rate, similar to the 2.52-fold difference observed.

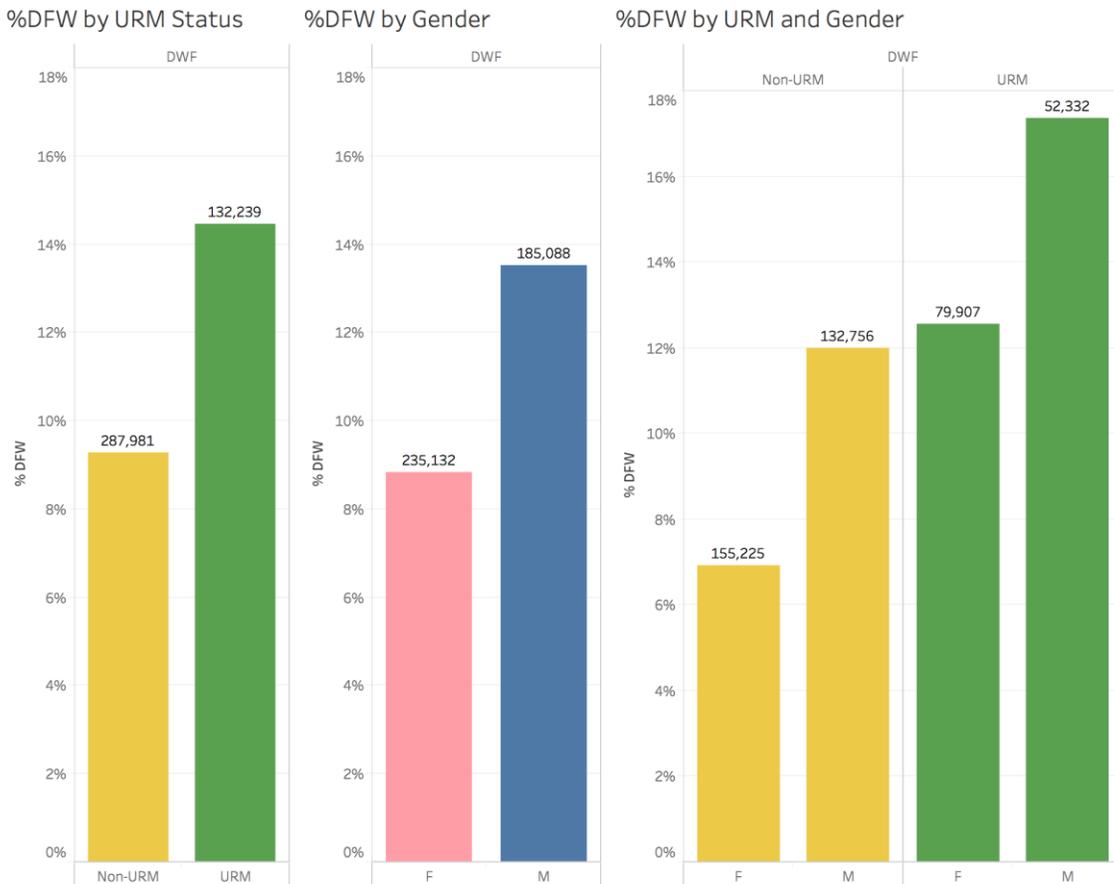


Figure 3. DFW Rates by URM and Gender Status.

A powerful factor affecting DFW rates is the student's high school preparation, as measured by the HS Eligibility Index, a combination of HS GPA and SAT (or ACT) scores. FTF were grouped by their HS Eligibility score and then the DFW rate was calculated for each group over all LD classes they had taken, Fig. 4. FTF with index scores between 2.6 and 2.8 have a 25% chance of getting a poor grade in a LD class, while FTF with index scores over 4.4 have less than a 2% risk of getting a poor grade. Thus, an 0.1 increase in index score gives a 1.1% decrease in the DFW rate. As the URM FTF's average HS Eligibility score is 0.15 less than for the non-URM FTF in this sample (3.5 vs 3.65), this difference explains only 1.6 percentage points of the 5.5 percentage point difference in LD DFW rates for URM and non-URMs (16.7% vs 11.2%). Thus only 30% of the differences between URM and non-URMs is explained by their lower HS eligibility scores.

Plotting DFW rates vs HS eligibility scores for URMs and non-URMs shows that even after correcting for index differences, there is a substantial gap of 4 to 5% between URMs and non-URMs, except at index scores above 4.0, where the gap goes away. There is only a .05 difference in average eligibility scores between males and females so little of the gender gap is explained by differences in HS eligibility scores. Unlike the URM plot, the gender gap does not narrow at high index scores. As a consequence, at high eligibility scores URM females do as well as non-URM females and much better than wither URM or non-URM males.

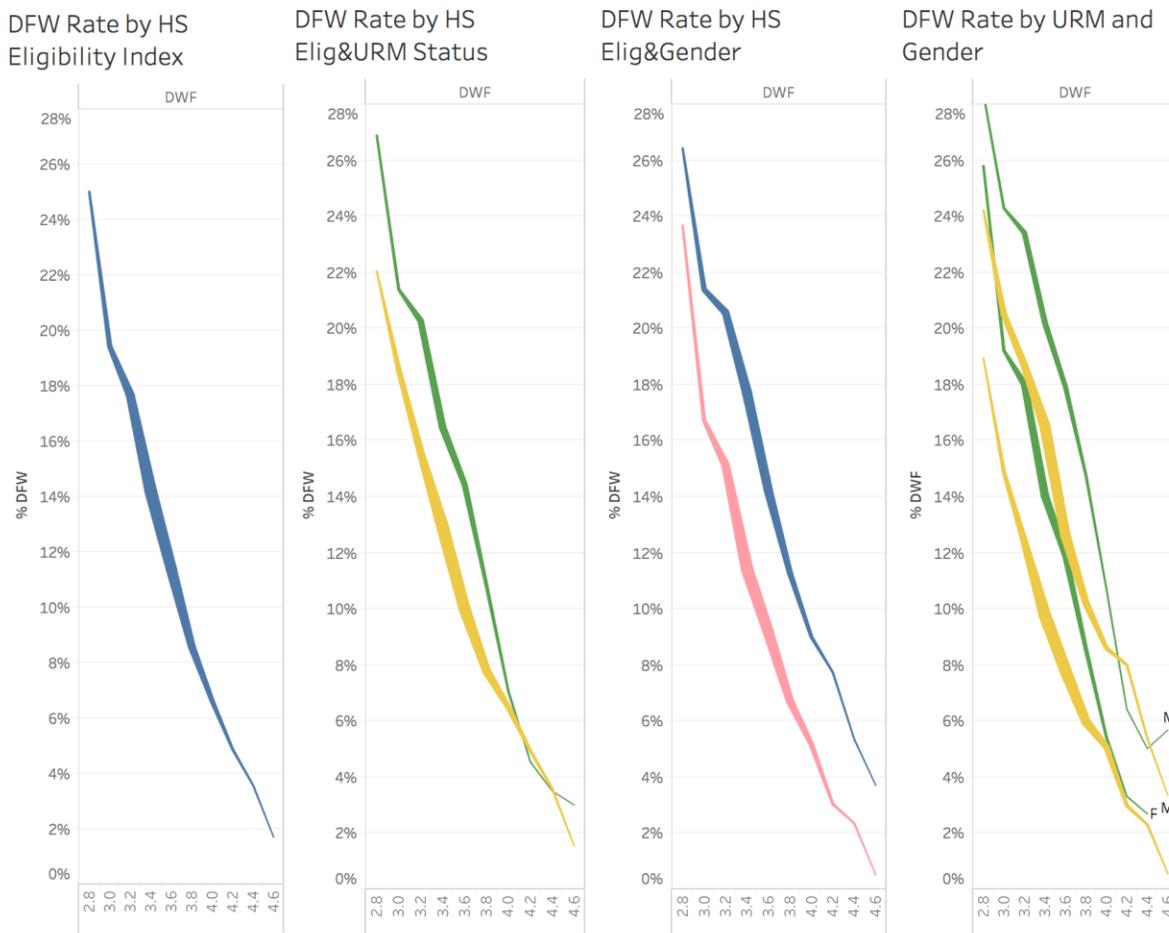


Figure 4. DFW Rates by HS Eligibility Index and URM and Gender Status for FTF in LD Classes. The HS Eligibility bins are labelled by the highest index score included, so the “2.8” bin includes all FTF with index scores from 2.61 to 2.8.

The DFW rates for General Education (GE), writing intensive and capstone (WI&Z) and majors/service (non-GE) classes at both LD and UD were compared for URMs and non-URMs in Fig. 5. In all types of classes URMs have significantly higher DFW rates (over 1.5-fold higher). Oddly, GE classes have lower DFW rates for LD classes (12.3%) than non-GE (14.5%), but higher rates for UD classes (8.7% vs 7.3% for non-GE). WI and Z classes have much lower DFW rates than other GE classes, 8.2% vs 12.6% at the LD level, and 6.3% vs 9% at the UD level.

## %DFW by Level&GE&URM

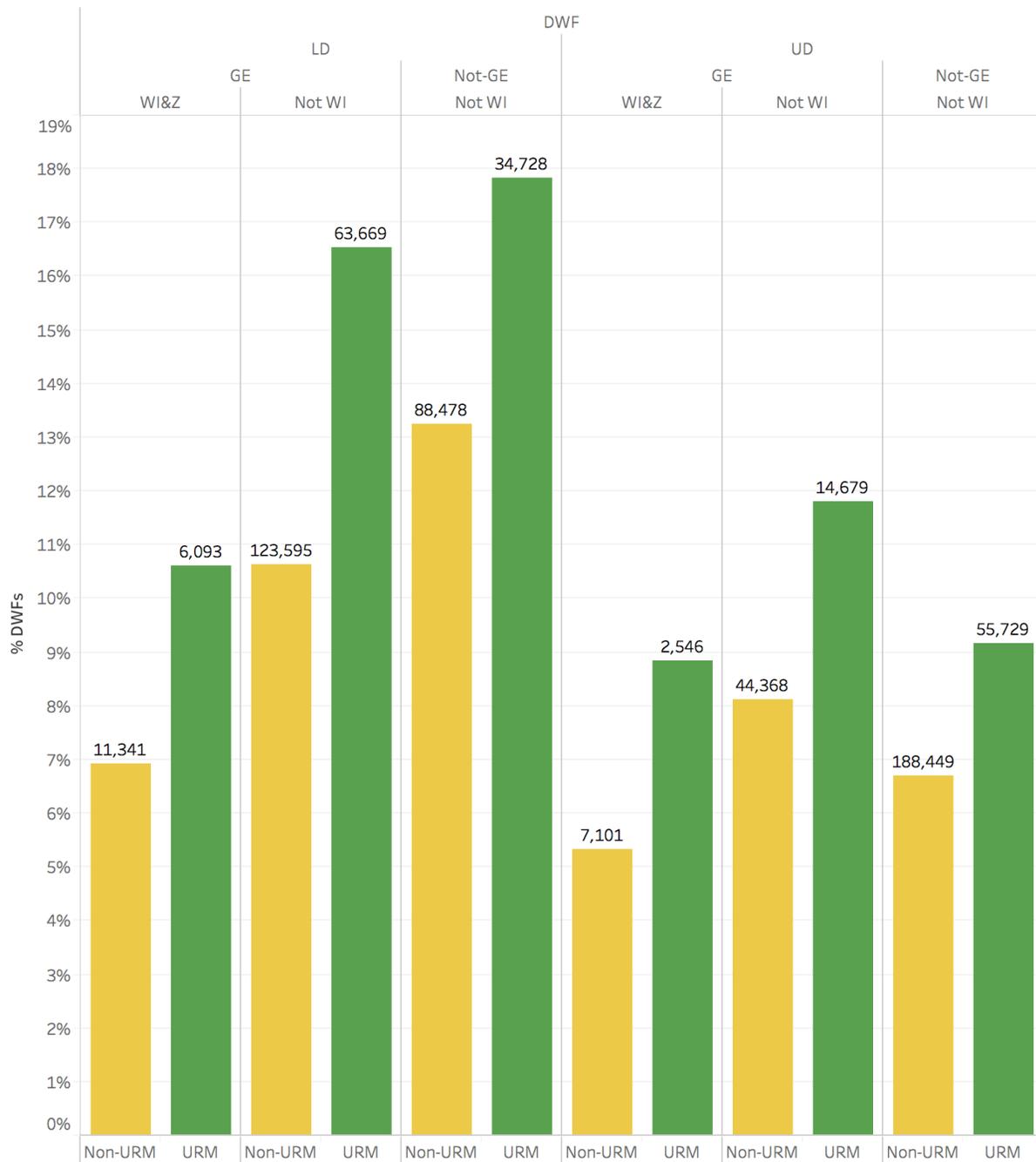


Figure 5. DFW Rates by Class Level, GE or non-GE and WI&Z for URMs and non-URMs. Numbers above the bars are the total sample size.

Unlike with URMs, there are differences in the DFW risk for males relative to females in the different class types, Fig. 6. In LD GE classes males are 40% more likely to get a DFW, but are more than 60% more likely to do so in a non-GE class. For UD GE classes males only have an 18% higher risk of a DFW, but have the normal 55% higher risk for non-GE classes. The reason for the male outperformance (or female underperformance) in GE classes is unknown.

### %DFW by Level&GE&Gender

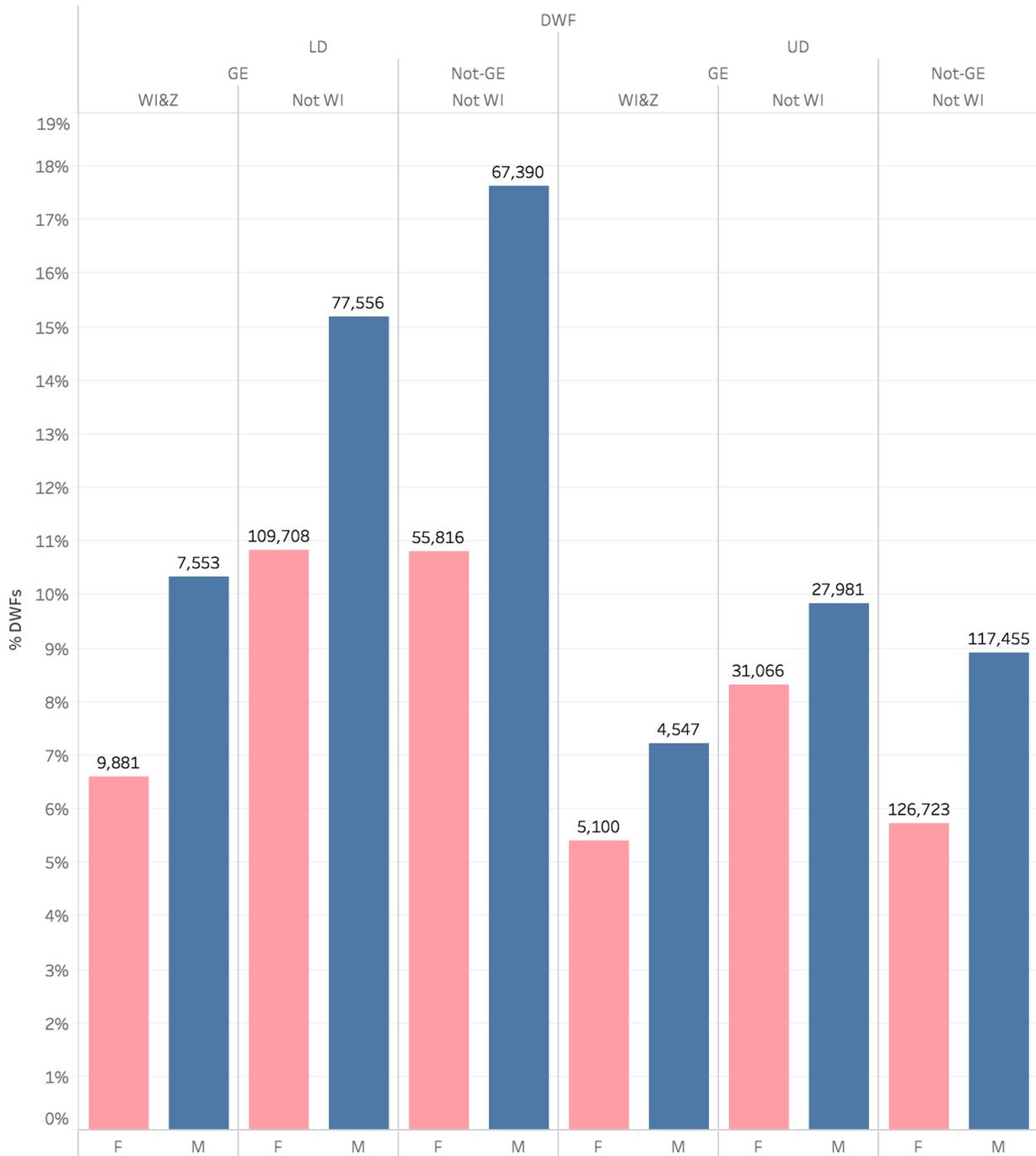


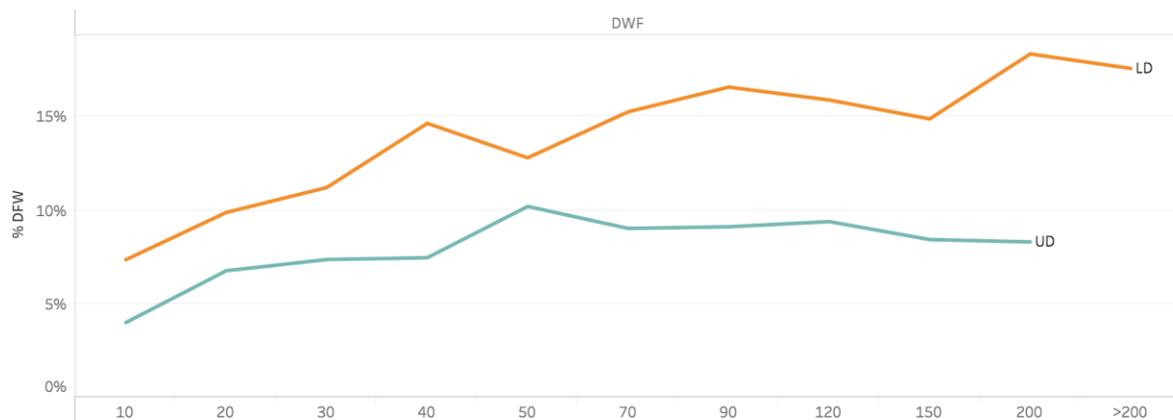
Figure 6. DFW Rates by Class Level, GE or non-GE and WI&Z for Females and Males. Numbers above the bars are the total sample size.

As UD classes tend to be smaller than LD, and the WI and Z classes have maximum caps of 30, the effect of lecture class size was examined in Fig. 7 to see if that explained the differences between LD and UD and WI and non-WI DFW rates (most labs and activities are part of mixed classes with lecture sections, so were not included in this analysis). Class size has a strong effect

on DFW rates with the rates more than doubling from 7.4% for LD classes of 10 or less to ~18% for LD classes of 200 or more and also more than doubling from 4% to 9% for UD courses. While the distribution of LD classes is skewed more towards large classes with many more classes of 120 or more and fewer classes of 40 or less, explaining part of the DFW gap between LD and UD, at all class sizes UD classes have DFW rates 30 to 50% less than the LD rate. This is especially pronounced for large classes, as the increase in DFW rate plateaus for UD classes at ~50, while it continues to increase for LD classes.

While the smaller class size for WI&Z explains part of their lower DFW rate, the DFW rate of 8.2% for LD WI&Z is well below the average for GE classes of 30, 11.2%, and the UD rate of 6.2% is also below the UD rate of 7.3%, thus there is some other grade advantage for WI&Z classes.

DFW by Class size



Class-size Dist-Level



Figure 7. DFW Rates by Lecture Class Size (upper figure) and % of Classes in a Size Range (lower figure) for LD and UD classes. After 50 the bins increase in size as the limited number of large class rooms leave some bins with no classes. The bump at 120, for instance, is due to there being two 120 seat classrooms.

Using just LD classes, the effect of class size on DFW rates by URM status and gender is shown in Fig. 8. The harmful effect of larger class size is much harder on URMs. For class sizes of 21 to 30 non-URMs have a DFW rate of 11.2% as opposed to the URM rate of 13.3%, but in classes of 200 or more the gap increases to 15.5% for non-URMs vs 24% for URMs, 1.5-fold higher. The gender gap, on the other hand, is much larger than the URM gap for small classes, but is much less in the larger classes. Thus, URM females are the group most harmed by larger class sizes, and non-URM males are the group least harmed. However, with DFW rates greater than 20% for all classes with enrollments greater then 40, URM males are still the group most at risk.

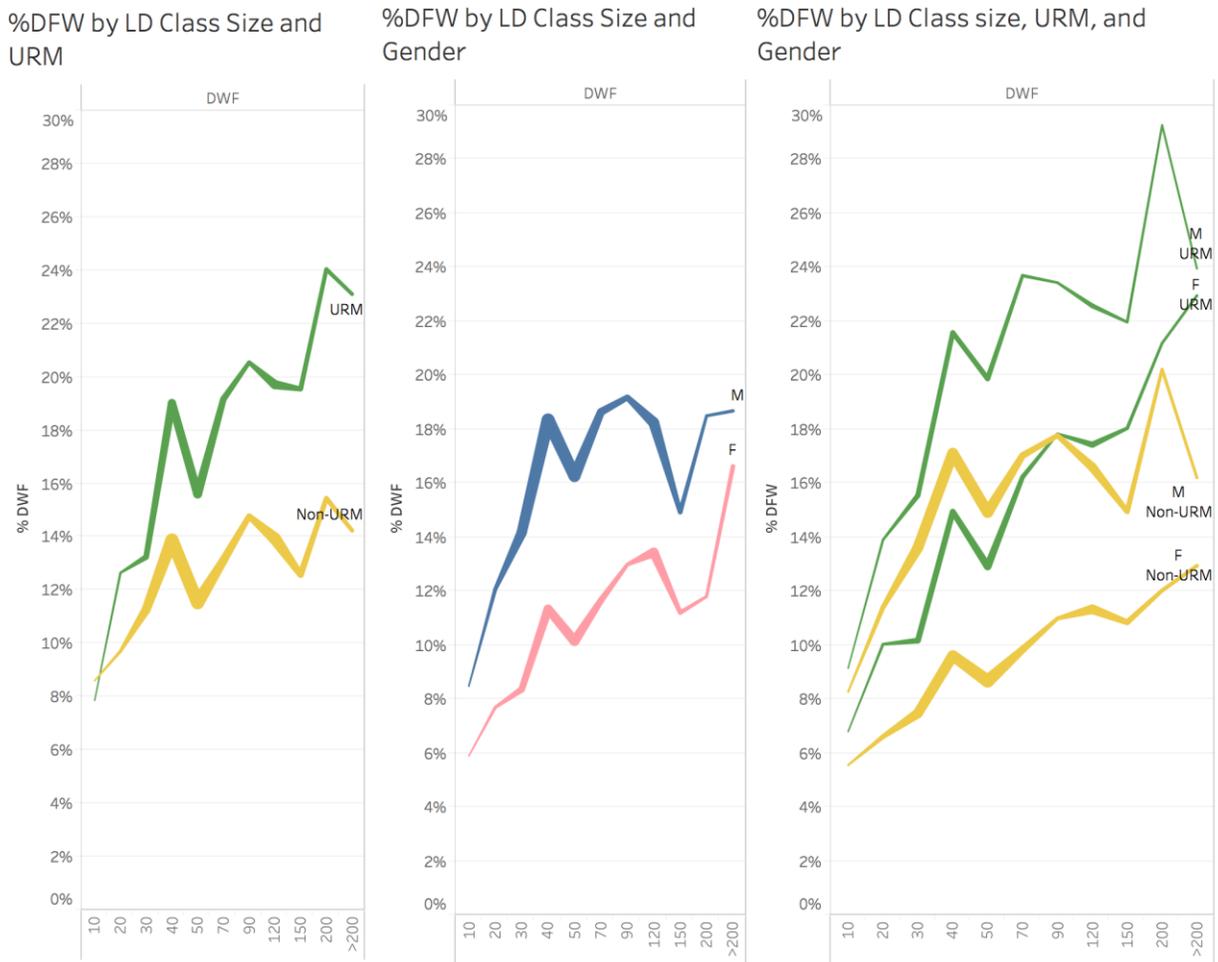


Figure 8. DFW Rates by LD Lecture Class Size and URM Status and Gender.

Another potential factor affecting DFW rates is subject. As the bump up in the LD DFW rate for class sizes in the 31 to 40 range corresponds to the most common MATH class size, the class size effect on LD DFW rates with and without MATH was plotted in Fig. 9. As surmised, removing MATH completely removed the bump in rate for the 31 to 40 LD class size, showing that there are also subject effects on DFW rates. The bump at class sizes of 150 to 200 is caused by Chemistry (data not shown). The cause of the bump at 70-90 is still unknown.

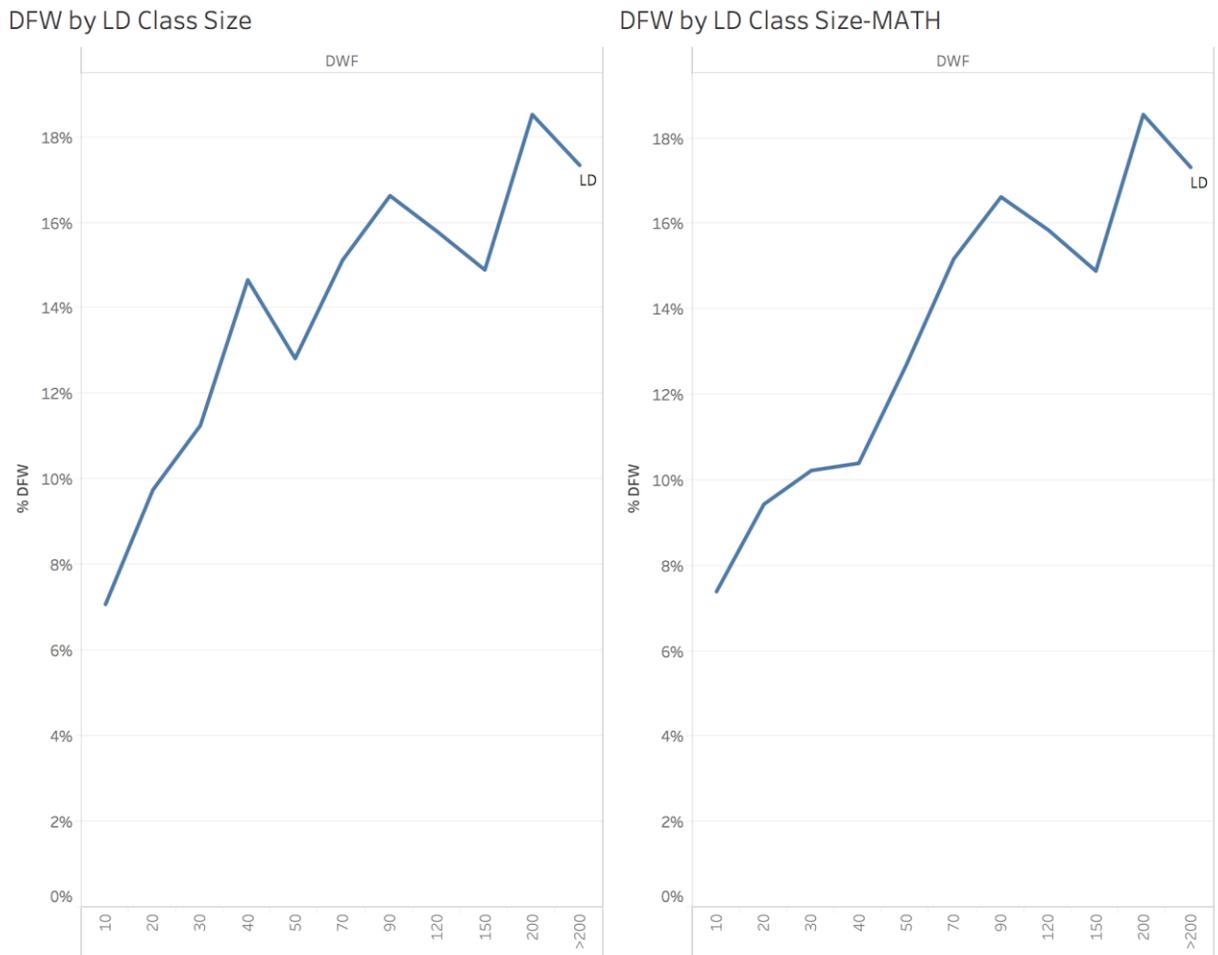


Figure 9. DFW Rates by LD Lecture Class Size with and without MATH.

As there appear to be some subject specific effects on DFW rates, the DFW rates and URM differences was calculated for all LD classes from each college, Fig. 10. At the College level DFW rates vary by over 3-fold, from a low of 6.7% for CME (4.7% for UED) to a high of 23.6% in Business. The URM gap also varies, though by much less, from a low of 1.28-fold higher in Business to a high of 1.57-fold higher in HFA. That the URM gap is so consistent despite the extreme variation in overall DFW rates suggests that the causes of the URM gap are general and not specific to a particular College. That Business has both the lowest URM gap and also the lowest percentage of URMs taking classes suggests that a lack of interest in a topic is not a cause of the URM gap.

%DFW for LD College Classes by URM

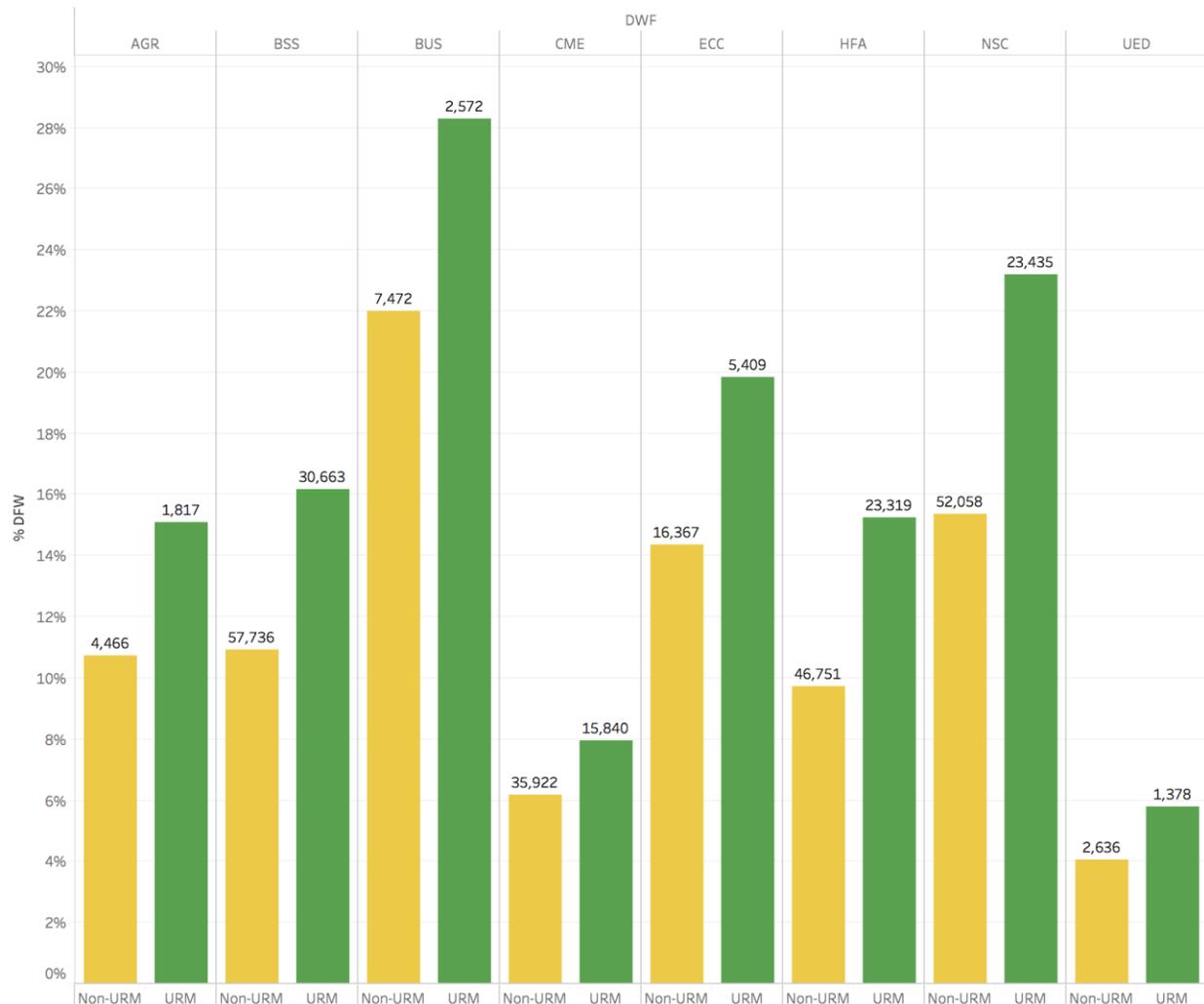


Figure 10. LD DFW Rates by College and by URM. Numbers above the bars are the total number of URM or non-URM taking LD classes in the college.

On the other hand, gender gaps vary more than the URM gaps, from a low of males having only a 1.09-fold higher risk of a DFW in Business to a high of 1.72-fold in CME, Fig. 11. That the two colleges with the lowest percentage of female students in their classes, Business and Engineering, also have the lowest gender gaps suggests that there may be some subject specific effects on the gender gap, though, again, these differences are not nearly as large as the variation in overall DFW rates.

%DFW for LD College Classes by Gender

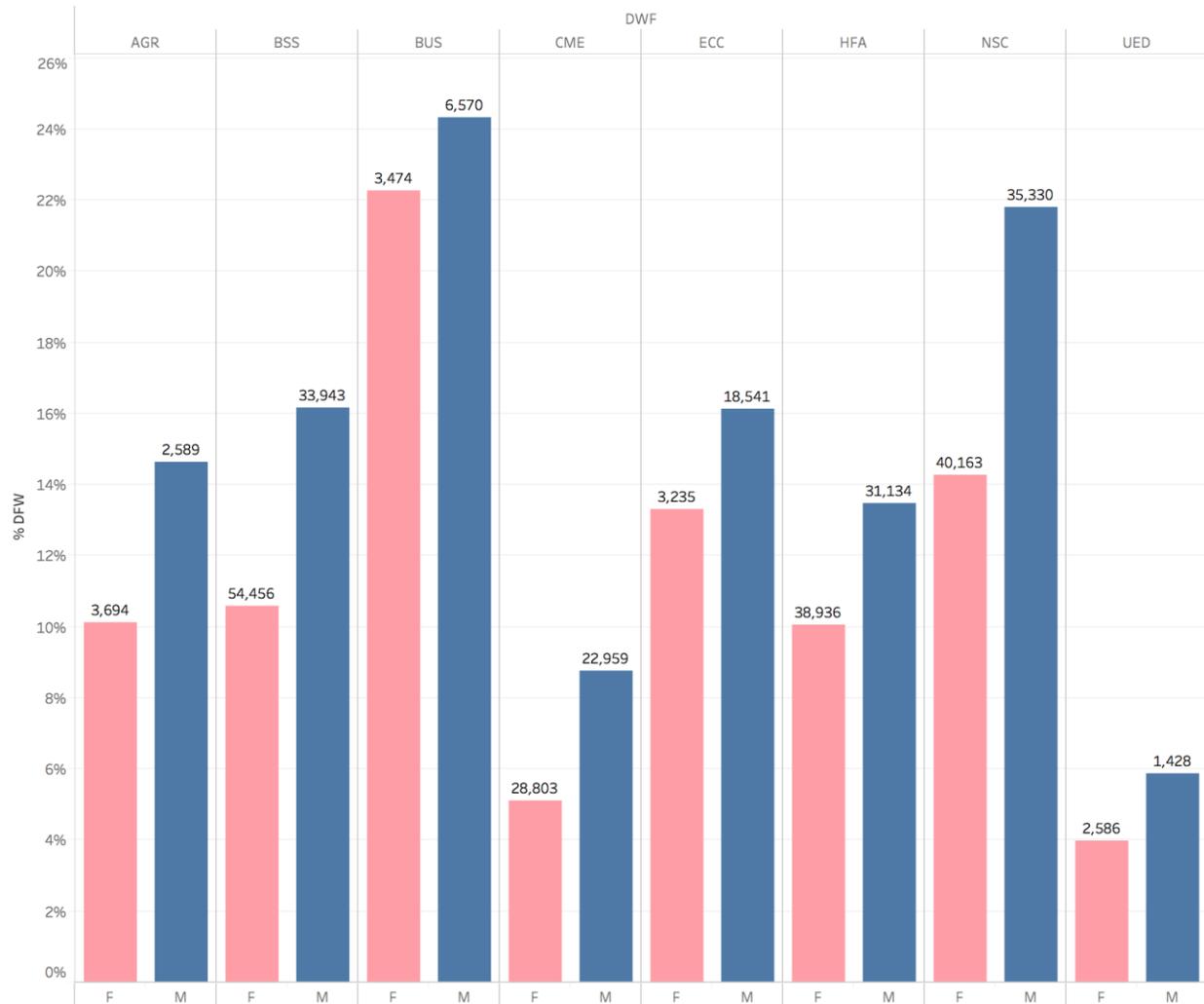


Figure 11. LD DFW Rates by College and by Gender. Numbers above the bars are the total number of males or females taking LD classes in the college.

Within a college DFW rates can vary over four-fold between departments. Fig. 12 shows the distribution of URM and non-URM rates for one college, BSS, chosen for its large number of departments and average overall DFW rate. Overall DFW rates vary from a high of 22.6% in Social Sciences to a low of 5.6% in Health and Community Services. This wide variation in DFW rates between departments within one college is puzzling.

Differences in DFW rates between URM and non-URMs vary from a low of no significant difference in Social Work to Sociology, where the URM DFW rate of 11% is 2.3 times the non-URM rate of 4.9%. Interestingly, both Social Work and Sociology have relatively high numbers of URM taking LD classes, so the size of the gap is not correlated with the popularity of the classes for URM. The spread is even larger when subjects from all colleges are used, ranging from highs of over a two-fold difference in Geoscience or Philosophy classes to near or below one in Spanish, Technology and University 105 (data not shown). Differences in the URM ratio between classes cannot explain the wide variation in overall DFW rates noted above.

BSS Coll-Dept-URM %DFW

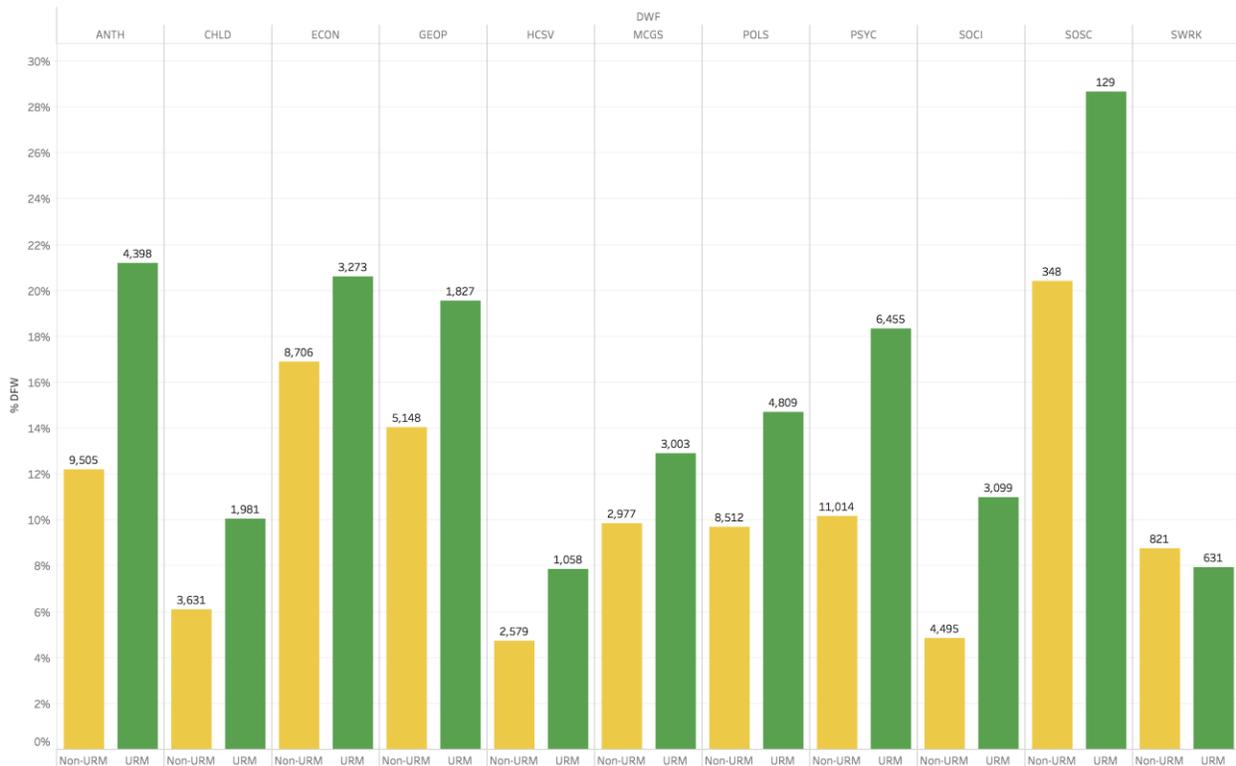


Figure 12. LD DFW Rates by Department and by URM for the College of BSS. Numbers above the bars are the total number of URM or non-URM taking LD classes in the Department.

The ratio of the DFW rate for males compared to females varies from a low of only 1.13-fold higher in Economics to 1.96-fold in Health and Community Services, Fig. 13. Economics has the highest percentage of males enrolled, 64% males and is also the BSS department where males do the best relative to females. The two departments with the lowest percentage of males, Child Development (11% male) and Health and Community Services (21% male) have the largest relative gaps in DFW rates, 1.89-fold and 1.96-fold, respectively. This suggests that gender preferences for classes is related to performance in those classes, unlike the case for URM status. While the gender factor explains a portion of the differences in DFW rate between departments, it is only a small factor and cannot explain the 4 and 5-fold differences in DFW rates between departments.

BSS Coll-Dept-Gender %DFW

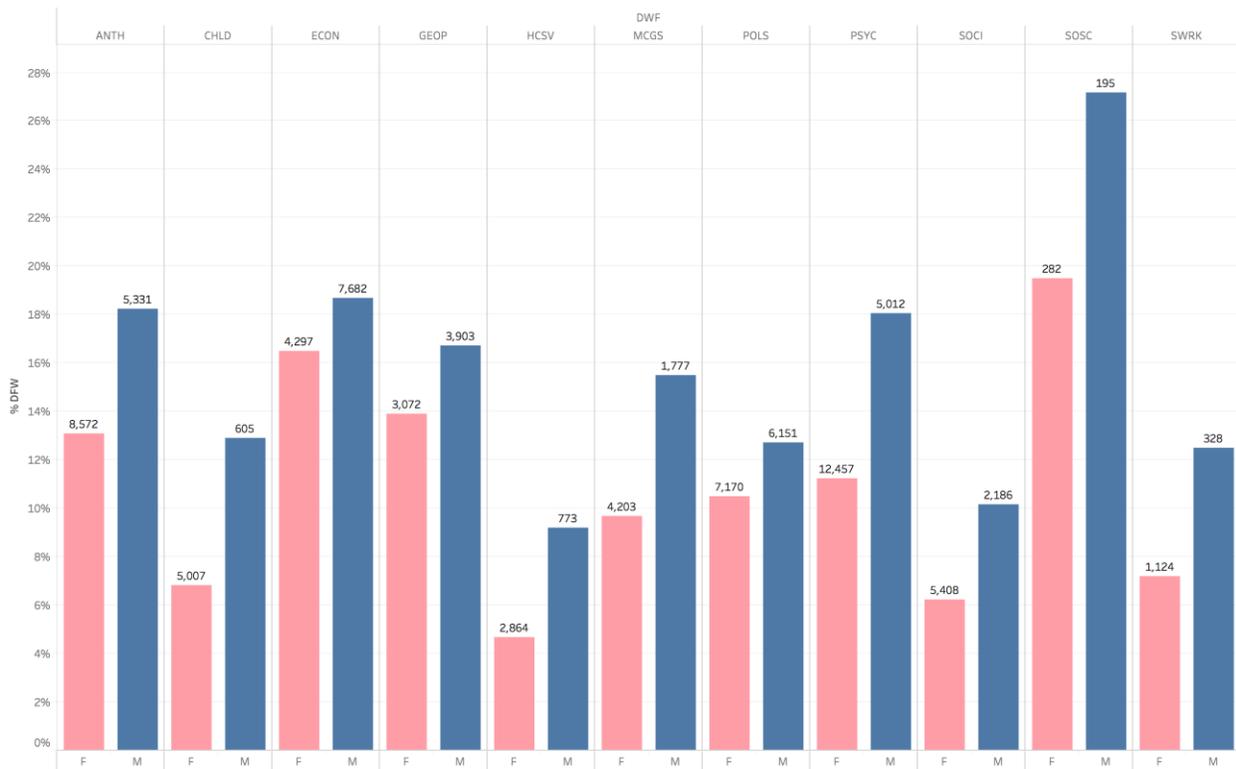


Figure 13. LD DFW Rates by Department and by Gender for the College of BSS. Numbers above the bars are the total number of males or females taking LD classes in the Department.



## Conclusions

This analysis of DFW rates shows that the same student factors correlated with differences in graduations rates, HS Eligibility score, URM status, and gender, all also strongly impact DFW rates. This makes sense as getting DFW grades leads to dismissal for poor grades, which then results in a failure to graduate. *Thus, reducing DFW rates and reducing the differences in those rates between URMs and non-URMs is key to meeting our graduation goals.* Fortunately, we can measure the effects of interventions on DFW rates much sooner and more easily than measuring graduation rates, as there is no need to wait many years, and the sample of sizes for grades are much larger than numbers of graduates as each graduate takes 40 or more classes.

Several general class characteristics have effects on DFW rates, from a two-fold difference between small classes and very large classes, a 1.7-fold difference between LD and UD, and smaller differences between GE and non-GE and between WI and regular GE classes. In general, the URM and gender effects are consistent across these different types of classes, though there are some unexpected effects, such as females, especially URM females, doing worse than expected in large enrollment classes, and males doing better than expected in GE classes.

The strongest class effect on DFW rates, though, is the variation in DFW rates between classes in different departments, and in the relative risk of URMs and males. Department DFW rates vary by over 10-fold and the relative risk for URMs varies from 0.6 to 2.3, almost a four-fold variation in the relative risk. Variation in the relative risk for males shows a similar pattern to URMs, though there is no connection between these two types of relative risk.

Determining why DFW rates and the relative risks for URMs and males varies so much from department to department will require examining DFW rates of individual classes and effect of the instructors of those classes on DFW rates, the subject of the next report, "Class and Instructor Effects on DFW rates."