Identifying Opportunities to Improve Systemic Inequity in Higher Education

Kameryn Denaro
Kimberly Dennin
Michael Dennin
Brian Sato
Working Paper #21-2
Identifying opportunities to improve systemic inequity in higher education

Kameryn Denaro*, Kimberly Dennin¹, Michael Dennin²³, Brian Sato¹⁴

1 Division of Teaching Excellence and Innovation, University of California, Irvine, California, United States of America
2 Division of Undergraduate Education, University of California, Irvine, California, United States of America
3 Department of Physics & Astronomy, University of California, Irvine, California, United States of America
4 Department of Molecular Biology and Biochemistry, University of California, Irvine, California, United States of America

* kdenaro@uci.edu

Abstract

It is well established that there is a national problem surrounding the equitable participation in and completion of STEM higher education programs. Persons excluded from science because of their ethnicity or race (PEERs) experience lower course performance, major retention, sense of belonging, and degree completion. It is unclear though how pervasive these issues are across an institution, from the individual instructor, course, major program, and discipline perspectives. Examining over six years of institutional data from a large-enrollment, research-intensive university, we present an analysis of opportunity gaps at the course section level between PEERs and non-PEERs to identify the consistency of these issues. From this analysis, we find that there is considerable variability as to whether a given course section taught by a single instructor does or does not exhibit opportunity gaps, although encouragingly we did identify that 34 percent of our instructor sample consistently observed no gaps. We also identified significant variation across courses within a major program, and found that certain STEM disciplines were much more likely to have courses that exhibited opportunity gaps relative to others. Across nearly all disciplines though, it is clear that these gaps are more pervasive in the lower division curriculum. This work highlights a means to identify the extent of inequity in STEM success across a university by leveraging institutional data. These findings also lay the groundwork for future studies that will enable the intentional design of STEM education reform by leveraging beneficial practices exhibited by more instructors and departments assigning equitable grades.

Introduction

A main consideration for educators and researchers examining means to improve higher education is the role that the institution plays in recreating and reinforcing structural inequalities. Related to this is how a student’s race and ethnicity differentially affects their experiences and outcomes in the education system. Identifying the root cause of this discrepancy is not a simple task, due to the complex nature of structural racism and the importance of intersectionality in regards to one’s identity and the effects of
The challenge for higher education institutions moving forward is how to address structural racism in the education system and to implement strategies to eliminate the racial opportunity gap. Evidence-based interventions that have been successful in these efforts include learning communities (7), distance education programs (3), undergraduate research and student organizations (5), light-touch values affirmation and utility-value interventions (6; 20; 21), use of alternative assessments (22), and the effectiveness of shifting to a diversity cognitive frame (4). Additionally, there has been increasing work on classroom practices and course structures that improve student outcomes, particularly for PEERs (2; 8; 10; 14; 23; 24). These include the incorporation of active learning pedagogies which emphasize peer to peer or student/instructor interaction, higher order problem solving, and facilitated, independent learning time. The addition of increased structure in these courses, through the incorporation of preparatory pre-lecture quizzes, weekly problem sets, and multiple timed assessments as opposed to the common one midterm/one final model, have shown to be beneficial for student learning as well. While these studies are promising, it is clear that adoption of evidence-based teaching practices has not occurred on an institutional level (2; 8; 10; 14; 23; 24).

Because of the variability in instructor practices, the undergraduate STEM
experience is likely to differ from course to course. This makes it likely that the previously described STEM equity issues [6,7,20,23] are manifested to varying degrees throughout a major program and across the institution. This paper seeks to establish a method to characterize the pervasiveness of equity issues across an institution and to identify STEM instructors whose courses show minimal grade penalties for PEERs. Follow up work can then identify the beneficial practices employed by these faculty and determine whether practices in one course have indirect impacts on student success in their other courses. Specifically, we aim to address the following research questions:

1. How can we identify specific STEM instructors that either do or do not consistently exhibit opportunity gaps in their course sections between non-PEER and PEER populations?

2. Using this methodology, are certain STEM departments more likely to exhibit opportunity gaps versus others?

3. Using this methodology, do opportunity gaps vary across an undergraduate major program?

**Materials and methods**

**Data**

Primary analyses were performed on a dataset containing 4,644 undergraduate courses at a selective research university in the Western US. Between the Fall 2013 and Winter 2020, we identified STEM courses that had at least 10 PEERs and 10 non-PEERs in the same course section. The inter-quartile range for % PEER in this sample was 28-49% PEERs. To control for class composition, we focus on these courses that had similar PEER representation.

The dataset included student demographics and transcript data including course enrollments and performance. Student demographic data included the student’s PEER status (Black, Latinx, Asian Pacific Islanders, and Native American students). Transcript data included information on each course (i.e. course number, subject, type, and level) and on student performance (i.e. grade in course). This study was approved by the institution's local Institutional Review Board (IRB #2018-4211) to study educational equity across the campus.

**Statistical Tests**

We used two different methods to explore whether there is a difference in grades given to PEERs and non-PEERs in a given course section. First, we conducted a Chi-square test of homogeneity to determine if there was a difference in the fraction of A’s and B’s awarded compared to C’s, D’s, and F’s. Second, we performed a two-sample t-test for independent samples (PEERs and non-PEERs) to test if there is a difference in the average grade received by the two groups. To test if opportunity gaps (between PEERs and non-PEERs) vary across an undergraduate major program (i.e. lower versus upper division), we used a two-sample t-test of the differences of AvgΔGP and AvgΔ%AB. For all analyses, we compared raw grades without incorporating students’ prior academic achievement, for reasons described in the discussion.
Results

Research Question 1: How can we identify specific STEM instructors that either do or do not consistently exhibit opportunity gaps in their course sections between non-PEER and PEER populations?

One measure of overall course performance for PEERs and non-PEERs can be determined by tabulating the numbers of A and B grades earned relative to C, D, and F grades for each population. Table 1 provides an example of this comparison for one course section (course section I) $\chi^2$-test results ($\chi^2 = 12.14, p = 0.002$). For this particular course section, 172 non-PEERs received either an A or B, whereas 67 PEERs received an A or B with the percentage of PEERs receiving A and B grades being significantly lower than the percentage of non-PEERs (46% of PEERs versus 63% of non-PEERs); the difference in the percent of A and B grades given to PEERs and non-PEERs ($\Delta\%AB$) was -17%. Furthermore, while PEERs made up 35% of Course Section I, the PEERs received only 28% of the A and B grades and received 44% of the C, D, and F grades.

Table 1. The observed and expected counts of grades versus PEERs status for Course Section I. The expected counts are in parentheses. The Chi-square test of homogeneity yields significant results $\chi^2 = 12.14, p = 0.002$.

<table>
<thead>
<tr>
<th></th>
<th>AB (expected)</th>
<th>CDF (expected)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-PEERs</td>
<td>172 (155)</td>
<td>100 (117)</td>
<td>272</td>
</tr>
<tr>
<td>PEERs</td>
<td>67 (84)</td>
<td>80 (63)</td>
<td>147</td>
</tr>
<tr>
<td>Total</td>
<td>239</td>
<td>180</td>
<td>419</td>
</tr>
</tbody>
</table>

Another means to identify the presence of an opportunity gap in a course is by calculating the grade point average (on a 4.0 scale) difference ($\Delta GP$) between PEERs and non-PEERs. Table 2 provides the summary statistics and $t$-test results for Course Section I. The difference in the average grade received by PEERs and non-PEERs was -0.39, a statistically significant difference ($t = -3.92, p < 0.001$). Similar to the grade distribution comparison, this test also provides evidence that PEERs received lower grades on average compared to non-PEERs (PEERs average = 2.41, non-PEERs average = 2.79).

Table 2. The mean and standard deviation of grades for Course Section I for non-PEERs and PEERs. The independent samples $t$-test yields significant results $t = -3.92, p < 0.001$.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-PEERs</td>
<td>2.79</td>
<td>0.95</td>
<td>272</td>
</tr>
<tr>
<td>PEERs</td>
<td>2.41</td>
<td>0.97</td>
<td>147</td>
</tr>
</tbody>
</table>

To determine how consistent these results were across multiple sections of the same course taught by a single instructor, we performed these same statistical analyses across five sections of the course previously presented in Tables 1 and 2. For this particular instructor, 3 of the 5 times the class was taught, there was a statistically significant opportunity gap present as determined both by comparing the distribution of A and B grades (Table 1), and the difference in GPA (Table 2) for PEERs and non-PEERs. The opportunity gap in GPA between PEERs and non-PEERs varied across the 5 years, but was consistently negative (average difference in grade points given to PEERs and
non-PEERs = AvgΔGP = −0.31). The gap in the percentage of PEERs receiving A and B grades and the percentage of non-PEERs receiving A and B grades also varied across the 5 years (AvgΔ%AB = −13%; see Table 5).

Table 3. The observed and expected counts of grades (A and B grades versus C, D, and F grades) versus PEERs status for Instructor J across 5 years. The expected counts are in parentheses.

<table>
<thead>
<tr>
<th>Year</th>
<th>PEERs</th>
<th>AB</th>
<th>CDF</th>
<th>Total</th>
<th>χ²</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No</td>
<td>131 (125)</td>
<td>91 (97)</td>
<td>222</td>
<td>2.36</td>
<td>0.1294</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>44 (50)</td>
<td>45 (39)</td>
<td>89</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>175</td>
<td>136</td>
<td>311</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>89 (89)</td>
<td>68 (68)</td>
<td>157</td>
<td>0.01</td>
<td>&gt; 0.999</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>37 (37)</td>
<td>29 (29)</td>
<td>66</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>126</td>
<td>97</td>
<td>223</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>172 (155)</td>
<td>100 (117)</td>
<td>272</td>
<td>12.14</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>67 (84)</td>
<td>80 (63)</td>
<td>147</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>239</td>
<td>180</td>
<td>419</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>No</td>
<td>167 (155)</td>
<td>94 (106)</td>
<td>261</td>
<td>7.38</td>
<td>0.0115</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>64 (76)</td>
<td>65 (53)</td>
<td>129</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>231</td>
<td>159</td>
<td>390</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>No</td>
<td>94 (81)</td>
<td>39 (52)</td>
<td>133</td>
<td>12.50</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>50 (63)</td>
<td>54 (41)</td>
<td>104</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>144</td>
<td>93</td>
<td>237</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. The mean and standard deviation of grades for Instructor J for non-PEERs and PEERs and the results of each of the independent samples t-test. The Grade Point Difference (ΔGP) between PEERs and non-PEERs is also displayed.

<table>
<thead>
<tr>
<th>Year</th>
<th>PEERs Mean (SD)</th>
<th>non-PEERs Mean (SD)</th>
<th>ΔGP</th>
<th>Test t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.48 (0.84)</td>
<td>2.70 (0.99)</td>
<td>-0.22</td>
<td>-1.93</td>
<td>0.0552</td>
</tr>
<tr>
<td>2</td>
<td>2.62 (0.87)</td>
<td>2.63 (0.97)</td>
<td>-0.01</td>
<td>-0.11</td>
<td>0.9152</td>
</tr>
<tr>
<td>3</td>
<td>2.41 (0.97)</td>
<td>2.79 (0.95)</td>
<td>-0.38</td>
<td>-3.93</td>
<td>0.0001</td>
</tr>
<tr>
<td>4</td>
<td>2.49 (0.94)</td>
<td>2.88 (0.95)</td>
<td>-0.39</td>
<td>-3.83</td>
<td>0.0002</td>
</tr>
<tr>
<td>5</td>
<td>2.42 (0.91)</td>
<td>2.95 (0.79)</td>
<td>-0.53</td>
<td>-4.68</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Table 5. The difference in the percentage of PEERs receiving A and B grades and the percentage of non-PEERs receiving A and B grades from Instructor J. The last column gives the difference in the percentage of PEERs who receive A and B grades and the percentage of non-PEERs who receive A and B grades (i.e. Δ%AB between PEERs and non-PEERs).

<table>
<thead>
<tr>
<th>Year</th>
<th>PEERs %AB</th>
<th>non-PEERs %AB</th>
<th>Δ%AB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>49%</td>
<td>59%</td>
<td>-10%</td>
</tr>
<tr>
<td>2</td>
<td>56%</td>
<td>57%</td>
<td>-1%</td>
</tr>
<tr>
<td>3</td>
<td>46%</td>
<td>63%</td>
<td>-17%</td>
</tr>
<tr>
<td>4</td>
<td>50%</td>
<td>64%</td>
<td>-14%</td>
</tr>
<tr>
<td>5</td>
<td>48%</td>
<td>71%</td>
<td>-23%</td>
</tr>
</tbody>
</table>

We now expand these analyses across all instructors within our dataset (n_J = 389).
Figure 1 displays the instructor Avg∆GP and Avg∆%AB between PEERs and non-PEERs. Each instructor represented in this plot taught a minimum of 4 sections of a single course. The median Avg∆GP across our instructors was −0.25, which equates to nearly a quarter point GPA drop for PEERs relative to non-PEERs. When looking at the Avg∆%AB, there is a considerable spread among our sample (median value of −9.41%). Using this methodology, we were able to identify 139 instructors who had zero course sections with a significant difference between grades earned by PEERs and non-PEERs, in other words, instructors whose course sections never exhibited an opportunity gap. Out of those 139 instructors, 45 were teaching lower division courses with the remaining teaching upper division courses.

Research Question 2: Are certain STEM departments more likely to exhibit opportunity gaps versus others?

Figure 2 provides an example of one department’s (Department K) Avg∆GP and Avg∆%AB. Typically, PEERs received between -0.44 and -0.13 grade points lower than non-PEERs in course sections taught within Department K. The median Avg∆%AB in this particular department was -12%; that is, in this department the percentage of PEERs receiving A and B grades was lower than the percentage of non-PEERs receiving A and B grades. For this particular department, there were 445 course sections included in the departmental analysis. Of those course sections, 36% showed a difference in the distribution of A and B grades compared to C, D, and F grades for PEERs and non-PEERs. 47% showed a difference in the GPA received by PEERs and non-PEERs and 36% of the course sections in Department K showed a difference in both statistical tests.

We then conducted this same analysis for all STEM departments within the study institution (nK = 33) to identify how consistent the opportunity gaps were across campus. Figure 3 displays the Avg∆GP of course sections taught within a department. The median Avg∆GP of STEM departments is −0.23 and the median Avg∆%AB of STEM departments is −9.47%, similar to our analysis looking at instructors across campus (Figure 1). The median percentage of a given department’s course sections with statistically significant opportunity gaps was 13%. We were able to identify 1 exemplar department which had no course sections with a difference in grades given to PEERs and non-PEERs. At the other extreme, there was one department identified that had differences in grades given to PEERs and non-PEERs in 61% of their course sections.

Research Question 3: Do opportunity gaps vary across an undergraduate major program?

While it is clear that there is considerable variation in the size of opportunity gaps in course sections taught across departments within a single institution, we were curious to see how variable these gaps were within a single major program. Figure 4 provides an example of the differences in average grades received by PEERs compared to non-PEERs for each course section taught in Department K disaggregated by course level (lower versus upper division). There was significantly larger gaps for lower division courses compared to upper division courses for Avg∆GP (t = −7.80, p < 0.001) and Avg∆%AB (t = −7.98, p < 0.001). While the overall Avg∆GP in Department K’s course sections was −0.28, this value was −0.36 for lower division course sections compared to −0.19 for upper division course sections. The difference in the percentage of PEERs and non-PEERs receiving A and B grades were larger for lower division courses compared to upper division courses (lower division Avg∆%AB = −15%, upper division Avg∆%AB = −7%).
Out of the 33 ($n_K$) STEM departments, we examined 20 ($n_K^*$) that taught a minimum of 10 lower division and 10 upper division course sections. The scatterplot in Figure 5 highlights that the lower division course sections within a department generally had larger opportunity gaps (i.e. wider/larger AvgΔGP and AvgΔ%AB) compared to upper division course sections. There were only a handful of departments where the opportunity gaps were higher in their upper division course sections compared to the lower division courses.

**Discussion**

In this study, we set out to shift the focus from macro level institutional data analysis to micro level data analysis at the course, instructor, and department levels, with the intention of gaining a better understanding of racial opportunity gaps and how they are distributed across the institution. In order to better understand the structures that contribute to opportunity gaps at the institutional level, we wanted to provide a method to identify instructors, course sections, individual STEM departments, and course levels within undergraduate STEM majors that are more or less likely to exhibit opportunity gaps. The results from this particular institution show that PEERs are on average receiving lower grades according to both $\chi^2$-test results and $t$-test results. Even though, 65% of instructors in the dataset have a gap in at least one of their courses, only 36% of the STEM course sections ($n = 4,644$) had a difference in performance between PEERs and non-PEERs, highlighting the variable presence of these gaps within a given instructor’s course. Our results show that the median percentage of a given instructor’s courses with gaps was 17% and the median percentage of courses within a given STEM department with an opportunity gap was 13%.

The distribution of gaps, existence of instructors without gaps, and low average of overall courses with gaps all point to the overlapping of structural influences of instructor, course level, and department that result in the average gaps observed at the institutional level. The clustering of gaps around specific course sections or disciplines versus a uniform distribution across the curriculum is consistent with the evidence of larger systemic issues that research has been documenting for years in higher education. Due to factors like systemic bias (12; 13; 15), unequal power structures (3; 5; 11; 13; 19), conflicting cultural models (5; 13; 15; 17), and the internalization of expectations (11; 12; 16; 20), PEERs are denied access to the same type of success as non-PEERS.

There are limits, however, to what these data can reveal about racial opportunity gaps at an institution. To begin with, we did not incorporate student demographics besides ethnicity, such as gender or first generation status, or prior academic achievement when examining course section opportunity gaps. It is common practice to control for various student characteristics when conducting similar analyses (5; 8; 19), yet there are also many instances of utilizing raw score when institutions report these data (5; 20; 25). We argue that not including student demographics and prior achievement in an analysis model more accurately captures the student experience, as undergraduates in a course do not consider their demographic characteristics or prior academic achievement when considering their course outcomes relative to their classmates. Similarly, future opportunities that reference a student’s academic performance, such as graduate school admissions or job interviews, also consider only the raw performance. Another limitation is that our analyses do not extend beyond student institutional data, and thus do not consider course structures, classroom practices, or instructor level characteristics. These are likely critical factors that impact the presence of an opportunity gap, so it is important the future work enables us to add these data to our models. And from an analytical perspective, since the distribution of grades at the course section level tends to be skewed, we used two different tests to
triangulate our results. This produces a conservative estimate of differences in opportunity between PEERs and non-PEERs. Thus, we likely have identified only the lower bound of opportunity gaps present across an institution.

By examining these data on an institutional level, we were able to identify interesting and perhaps unexpected patterns. One is the finding that the presence or absence of opportunity gaps within a given course for a given instructor were not consistent. The majority of instructors in our dataset had some course sections where an opportunity gap was present while other instructors where it was not. This perhaps highlights the variability in instructor practices. While a considerable amount of empirical data does not exist to support the idea of variability in implementation of a single instructor’s practices, Owens et al. (26) demonstrated using classroom audio analyzed by the Decibel Analysis for Research in Teaching (DART) algorithm that a course section varies considerably from lecture to lecture in regard to the amount of active learning implemented. It is not a stretch to assume that examining a given lecture periods across distinct offerings of a course would result in variability as well. While core structural and pedagogical approaches may remain the same for faculty, their presentation, degree of buy-in sought from students, and communication around course policies can still among iterations of a given course.

Another reason for variability in the instructor course sections may be due to varying demographics of the course staff and students. For example, having a PEER teaching assistant may increase belonging and decrease imposter syndrome. From the student population perspective, while student demographic characteristics in our dataset were fairly consistent from section to section, differences may exist that do not manifest themselves as easily as ethnicity or gender, including non-cognitive factors like sense of belonging or growth mindset (5–7; 11–13; 15; 17). And while we focused on the percentage of PEERs in a course section, the total number of PEERs in a course may create a higher sense of belonging (for example 2 PEERs in a class of 10 versus 20 in a class of 100), so course enrollment may play a role as well (11; 12; 25).

We also found that the prevalence of opportunity gaps within an institution is not consistent across STEM programs. This is reflected in prior work. For example, Yang and Barth looked at the distribution of male and female students in STEM majors at two different public universities, one in the Southeast and the other in the Midwest. An analysis of 1,848 students showed that biology tends to have a larger representation of women when compared to other STEM majors like computer science, engineering, mathematics, and the physical sciences including physics, geology, and chemistry (27). In another study, Riegle-Crumb and King utilized data from the Educational Longitudinal Study (ELS) of 2002 to follow a cohort of approximately 15,000 students who were enrolled in a 4-year degree-granting institution in 2006. They found that Black males make up 28.5% of physical science/engineering majors while Black females make up only 7.3% of physical science/engineering majors. Comparatively, biology is much more gender balanced, with Black males making up 7.0% of majors and Black females making up 8.6%. This trend holds true for white and Hispanic males and females (28).

And perhaps not surprisingly, our data revealed that lower division courses within a given STEM major were more likely to exhibit opportunity gaps relative to upper division courses. This is supported by considerable literature highlighting the gatekeeper nature of lower division courses (5; 11–13; 15) and may also reflect that the disproportionate loss of PEERs in these courses leads to a more homogeneous upper division undergraduate population in terms of prior academic achievement. But it may also be a factor of differing grading practices in courses of these two levels. The lower division courses are more likely to use norm-referenced grading, where regardless of overall achievement, a fraction of the course section is destined to earn a C or lower (29–31). From the instructor and institution perspective, we would argue that there is
not a pedagogically-beneficial reason to utilize distinct course policies for lower and upper division courses, and plan to examine whether these are contributing to these differing opportunity gap outcomes.

In the rapidly changing higher education landscape, the increased use of data in higher education to evaluate successes and failures enables administrators to make decisions in an empirical fashion. Our analysis can be replicated within one’s own institutional context to show the persistence of racial opportunity gaps, but it is important to note that it does not provide an explanation as to what is causing these gaps. In our particular dataset, we highlighted exemplar instructors and departments, with 139 instructors and 1 department without a significant Avg∆GP in their course sections. By providing a way to identify these exemplar instructors and departments, a more detailed, qualitative examination can help to better identify the practices and policies that result in these more equitable outcomes.

Acknowledgments

We would like to thank the individuals involved in facilitating the acquisition of the data utilized in our analyses. We also thank Elizabeth Park, Emily Penner, Chandlyr Denaro, Veronika Rozhenkova, and Hal Stern for critical discussions and feedback. This work was supported by the National Science Foundation (NSF DUE 1821724).

References


Figures

Fig 1. Histograms of the opportunity gaps between PEERs and non-PEERs per course section at the instructor level. Histogram of the average grade point difference (AvgΔGP) and the average difference in the percent of A and B grades given to PEERs and non-PEERs (AvgΔ%AB) of a given instructor’s course sections. The red line represents the median on each plot.
Fig 2. Boxplot of AvgΔGP and AvgΔ%AB per course section for Department K.
Fig 3. Histograms of the differences in grades (AvgΔGP and AvgΔ%AB) given to PEERs and non-PEERs per course section at the department level. The red line represents the median on each plot.
Fig 4. Boxplot of $\text{Avg}_\Delta \text{GP}$ and $\text{Avg}_\Delta \%\text{AB}$ for Lower and Upper Division Course Sections in Department $K$. Within each plot the lower division course sections are displayed on the left and upper division course sections are displayed on the right.
Fig 5. Scatterplots of the differences in grades given to PEERs and non-PEERs in lower and upper division courses at the department level. The plot on the left is for department AvgΔGP and on the right is for department AvgΔ%AB. The red line ($y = x$) represents similarity in grades for lower and upper division courses (points above the line represent departments that have larger opportunity gaps in lower division course sections).