A Statistical Analysis of the Informed Placement Process (IPP) at UCLA

Final Project Team 1: Tishi Avvaru, Mukhil Guna, Michael Gureghian, Naomi Suzuki, Sydney Tomsick

June 7, 2025

Abstract

This study investigates how student self-assessments and demographic characteristics influence placement outcomes in UCLA's English composition courses through the Informed Placement Process (IPP). Using data from 980 students including IPP writing scores (based on performance in a writing task), final class placement levels, post-IPP survey scores, course preference, and demographic information, we created a binary variable, sec_grader, to indicate whether a student's writing task required a second evaluator due to a mismatch between their course preference and their writing score. Logistic and linear regression models were used to examine (1) the relationship between survey scores, demographic data, and the need for a second rater, and (2) how the involvement of a second rater relates to students' final course placements.

Findings show that higher post-IPP survey scores, which correlate with greater student confidence in their writing ability, significantly reduce the likelihood of requiring a second rater. Demographic factors, including underrepresented minority (URM) status, first-generation status, Pell Grant status, residency status, and gender, did not significantly predict the need for a second rater. However, URM status was a significant predictor of lower final placement levels. Further analysis revealed that international students and students requiring a second rater had lower post-IPP survey scores, while students whose parents had completed college scored higher on the survey. These results suggest that structural inequities may influence final course placement outcomes and confidence in one's writing abilities. Limitations include potential rater bias and the inability to evaluate long-term academic performance. Future work should consider student outcomes post-placement to further assess the fairness and efficacy of the IPP.

1 Introduction

This project uses data obtained from UCLA's English Department regarding student placement into English Composition courses (1) ranging from Eng Comp 1 - 3 based on student course preference, performance in a writing task (IPP score), and a post-IPP survey. UCLA's Informed Placement Process relates to the institution's Entry Level Writing Requirement, which can be fulfilled by completing IPP, demonstrating proficiency through test scores (e.x. AP Lang or AP Lit), or UC-transferable college English composition courses. The data is anonymized, both in regards to students and raters, and shows a student's final placement, course preference, IPP score, post-IPP survey, and demographic information.

2 Background

There are 1028 observations in this dataset (N = 1028) and 10 variables: Student UID (anonymized to show only the last four digits, not used in this report), Final ENG COMP Placement (Final Score), Student's Course Preference, IPP Score 1, Post-IPP Survey, urm, first_gen_bachelors, ever_pell_fl, residency_apb, and gender. The main response variable being used in this report is Final ENG COMP Placement (Final Score). To answer both parts of our research question, we will be using Student Course Preference, IPP Score 1, Post-IPP Survey (which correlates to a students' confidence in their writing ability, rated 16 - 48), urm (if the student is an underrepresented minority, classified as students who identify as African American, Hispanic/Latino, or American Indian/Alaskan Native), first_gen_bachelors (if the student is first generation), ever_pell_fl (if the student has ever been awarded a Pell Grant), residency_apb (the student's residency status within the United States), and gender.

We used a created variable in our analysis: sec_grader. This variable measured whether or not the students IPP score and course preference matched, with 1 representing a mismatch. These mismatches indicate if the student required a second opinion. The variable sec_grader is 1 if the student required a second rater, and 0 if not (meaning the IPP score and student course preference matched).

3 Exploratory Data Analysis

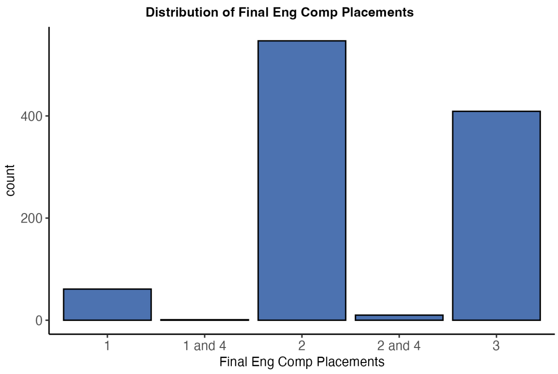
Table 1 below shows the amount of NAs found in our dataset. Overall, we saw very few NAs, with no category being over 2.5 percent composed of NA values.

Variable	Missing_Count	Missing_Percent
Final.ENG.COMP.Placement..Final.Score.	0	0
Student.s.Course.Preference	8	0.79
IPP.Score.1	9	0.88
Post.IPP.Survey	17	1.65
urm	18	1.75
first_gen_bachelors	18	1.75
ever_pell_fl	18	1.75
residency_apb	18	1.75
gender	23	2.24

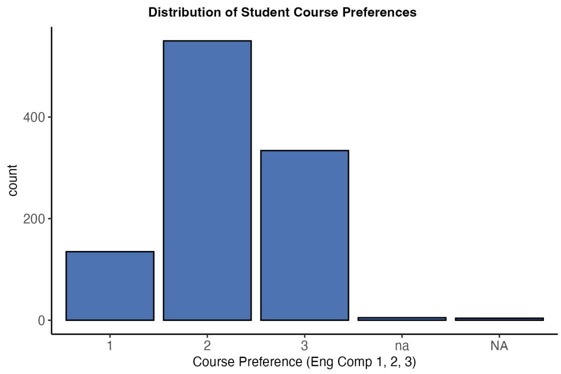
Table 1: Distribution of missing values

Figures 1, 2, and 3 below show histograms depicting the distributions of all of our variables. Figure 1 includes the distributions of the variables final placement score, course preference, IPP score, and post IPP survey. For course preference, we can see that most students rank themselves at a 2, the middle of the three courses. More students rank themselves at a 3 than a 1. This correlates to the final placement score and IPP score variables, which show that most students score at and get placed into level 2, then 3, with level 1 having the lowest placement and scoring numbers. Lastly, the plot shows that the post IPP survey variable is mostly normal, with a negative skew. Many students had scores in the high 30s and low 40s, demonstrating that they were moderately confident in their writing skills (the maximum score is 48).

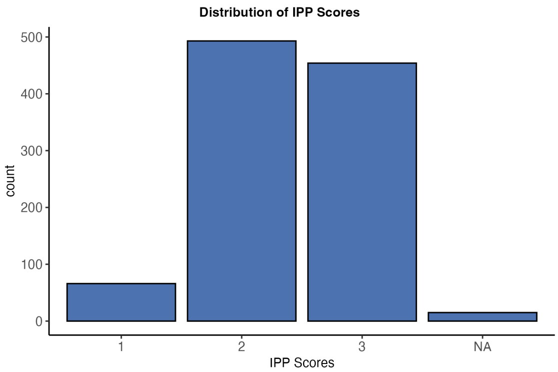
One issue we immediately noticed is that while most observations in the final placement score and IPP score variables took the form of a single value from 1 to 3, some had an “E”, “and 4”, question mark, or “high”/“low” classifier attached to the number (e.g. “2E”, “2 and 4”, “3?”, “high 2”). For the IPP score distribution plot, we removed these extraneous classifications. We will discuss how we remedied this issue more in the ”Data Cleaning” section.



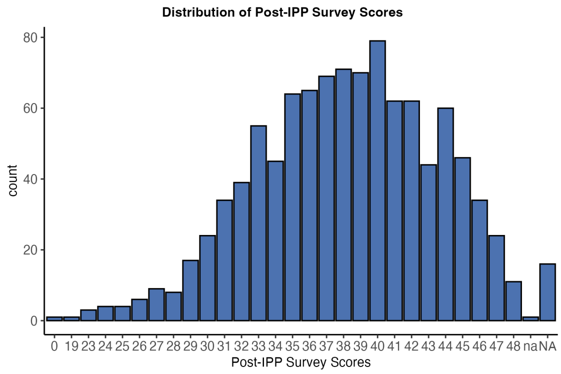
(a) Final placement score



(b) Course preference



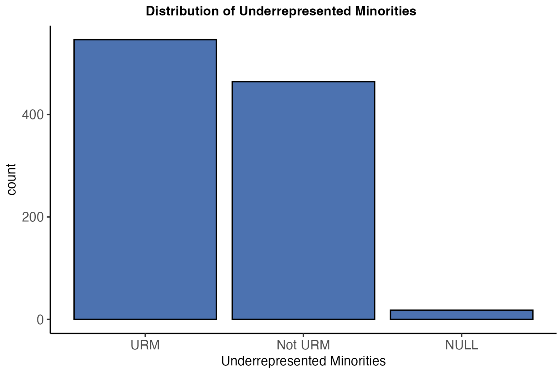
(c) IPP score from reading/writing task



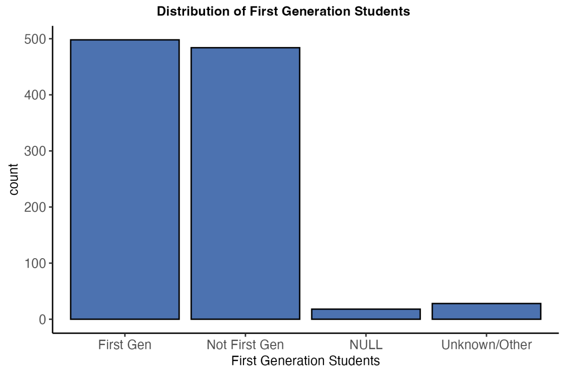
(d) Post IPP Survey

Figure 1: Distributions of Placement Variables

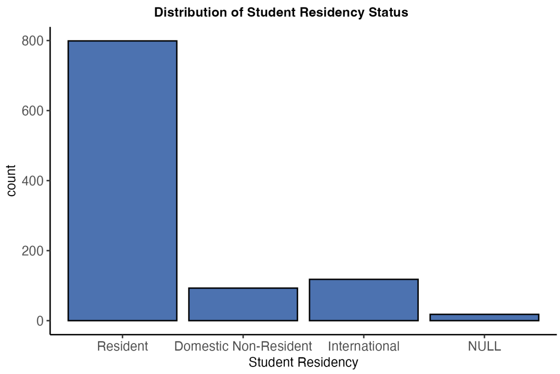
Figure 2 below shows the distributions of our demographic variables. There is generally an even split when looking at the distribution of URMs, students who have and haven’t received a pell grant, and first generation students. However, for the residency variable, most students are US residents while international and non-resident students are a small portion of the population.



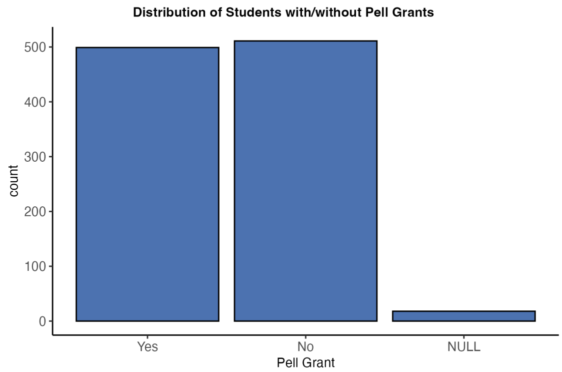
(a) Underrepresented minority status



(b) First generation status



(c) Residency status



(d) Pell grant status

Figure 2: Distributions of Demographic Variables

Another issue we found through our distribution plots is that there are missing values that are classified as something other than NA or NULL. For example, we see that in the student course preference and post IPP survey variables, there are some missing values classified as NA, and others as the character value "na". Further, in the first-generation variable, there are some missing values classified as NULL, and others as the character value "Unknown/Other". This is important to be aware of so that we can eliminate all missing values in data cleaning.

Lastly, figure 3 below depicts the distribution of our created variable, sec_grader, showing that there were more students who did not require a second rater.

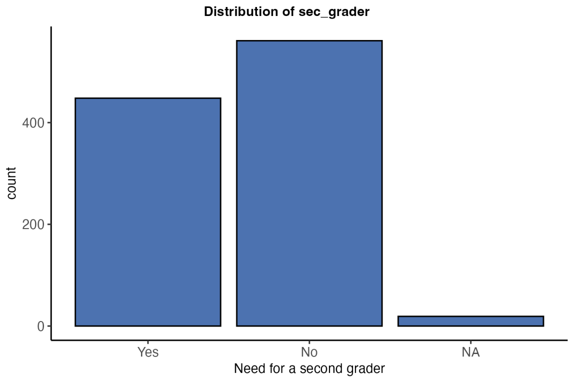


Figure 3: Distribution of the variable sec_grader

4 Data Cleaning

Before modeling, we cleaned our data by eliminating NA values, transforming the variables into the correct class, and standardizing the IPP score columns.

As we discussed in the previous section, the final score and IPP score variables had observations with extra classifiers atached to the course level number (e.g. "2E", "2 and 4", "3?", "high 2"). The "E" identifies English language learners and the "and 4" identifies multilingual students who are required to take a mentoring seminar along with their main IPP level class. In cleaning, we eliminated all extraneous classifiers, making each observation in the final score and IPP score variables only contain a single value from 1 to 3 representing the IPP class level.

Further, the post IPP survey variable is meant to only have values ranging from 16-48, so we eliminated the values that were outside that range. Laslty, we transformed all 4 of the placement variables (final score, course preference, IPP score, and post-IPP survey) to be numerical variables since they were originally stored as categorical variables.

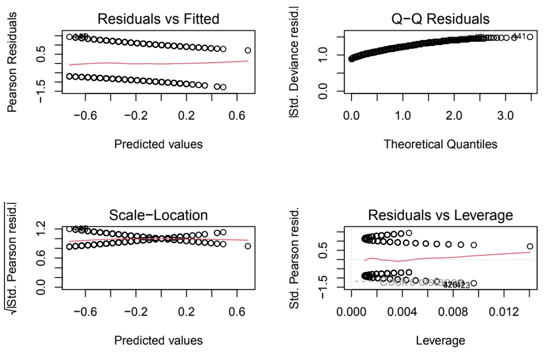
Lasly, we eliminated the NA values. 18 observations had missing values for every demographic variable (urm, first generation, pell grant, residency, and gender). Since these observations did not contain any information about the students' demographics, they were not helpful to answer our research questions and we removed them from our dataset. We also removed the NA values for the placement variables. The missing values for the student course preference and post-IPP survey variables were due to students neglecting to submit them. The data provider was unsure why there were missing values in the IPP score variable. Since these 3 variables contained missing values in under 2 percent of the data, and because we either weren't sure why the missing values existed or they were random, we eliminated the rows with missing values.

Our cleaned dataset contains 980 observations and 9 variables (we eliminated the UID variable). This is only 48 observations less than our original dataset given the small number of NA values.

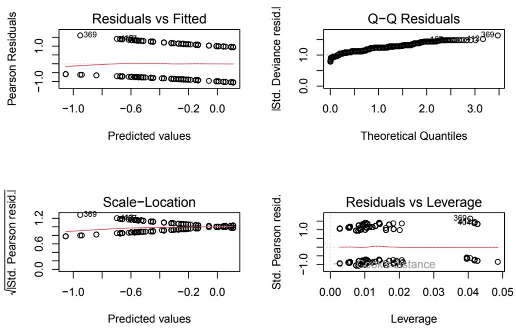
5 Methods and Results

To address our first point of analysis — comparing student student survey scores with their need for a 2nd rater — we utilized three logistic regression models. For our second point of analysis — looking at how students' final placement into Eng Comp 1, 2, or 3 are impacted by whether or not their case required a 2nd rater — we used two linear regression models. Below we describe and analyze the results of each model in detail, starting with the three logistic regression models.

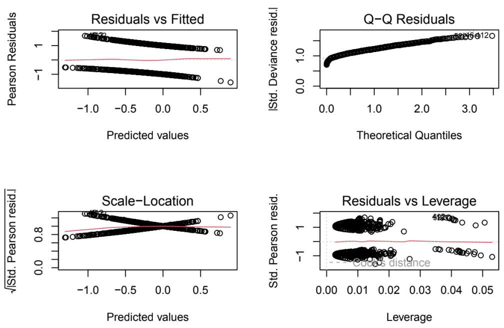
The diagnostic plots for all of our logistic models are passable, meaning that our models are able to be used effectively to make observations. The diagnostic plots for the third model are worse than the plots for models 1 and 2, but we find this acceptable, as model 3 only reiterates what we find in models 1 and 2.



(a) Diagnostic plot for model 1



(b) Diagnostic plot for model 2



(c) Diagnostic plot for model 3

Model 1 evaluates how the post-IPP survey score alone predicts whether a student required a second rater. Figure 5 shows that the coefficient for the post-IPP survey is -0.048, indicating a negative association. This means that for each one-unit increase in a student's survey score, the log-odds of needing a second rater decrease by approximately 0.048 units. This coefficient is statistically significant at a = 0.01, confirming that **students with greater confidence in their writing (reflected by higher survey scores) were less likely to need a second opinion.** However, the coefficient -0.048 is very small, meaning that while confidence does influence the chance of requiring a second rater, the effect is gradual, not dramatic.

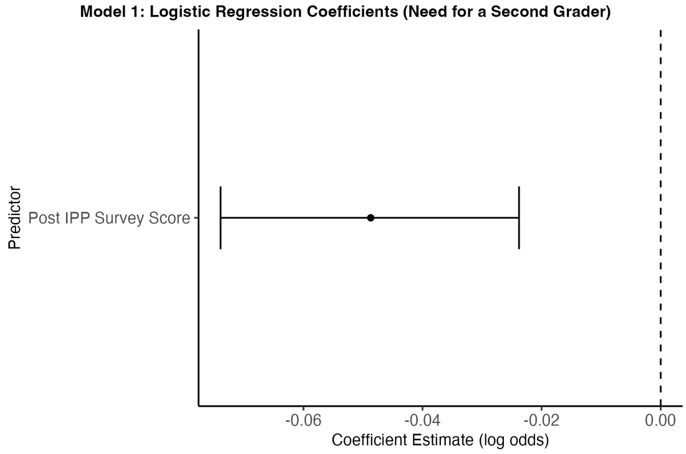


Figure 5: Model 1 Logistic Regression Coefficients

Our second model explores if demographics predict a students need for a 2nd rater. As we see in figure 6, only under-represented minority status (urm) looks be slightly significant predictor at 0.0662, but it is not significant if our a = 0.01, or even 0.05. As such, we see that **demographic factors do not impact a student's need to be rated by a 2nd rater.** The graph supports this, as the confidence intervals for all variables cross zero, visually indicating that these predictors do not have meaningful individual effects.

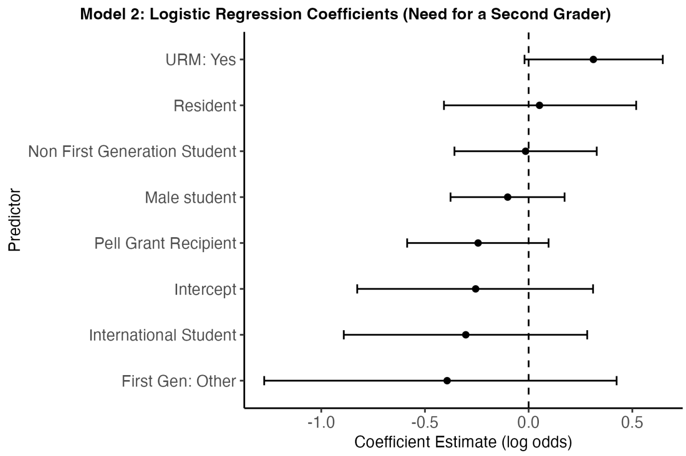


Figure 6: Model 2 Logistic Regression Coefficients

Our third model combines the analysis done in models 1 and 2, considering how demographics and survey scores impact the need for a 2nd rater. This model carries the same conclusions, showing that **the post-IPP survey is the only significant variable when considering sec_grader**

at a = 0.01.. All demographic variables remain insignificant, and their coefficients are relatively small and unstable, reflected by their wide confidence intervals in figure 7. Together, our three logistic regression models suggest that student confidence, rather than demographic background, is the primary driver of whether a second rater is needed.

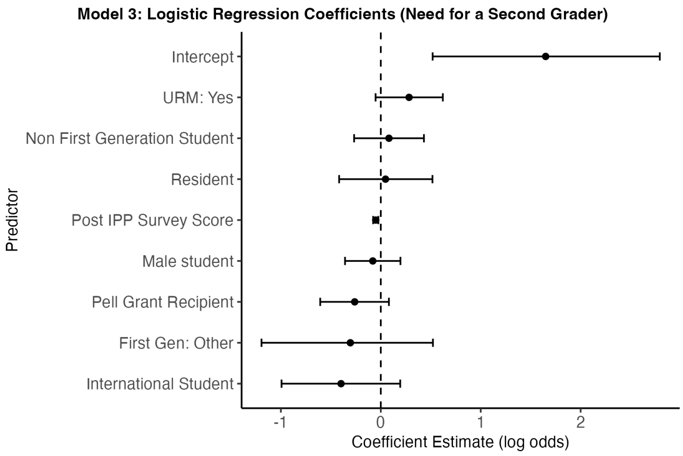
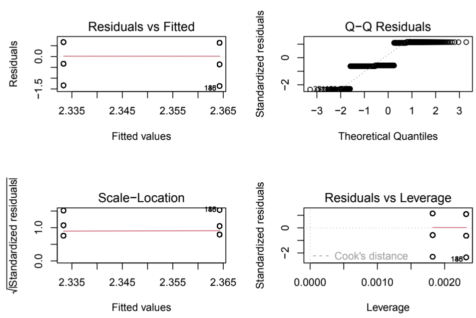
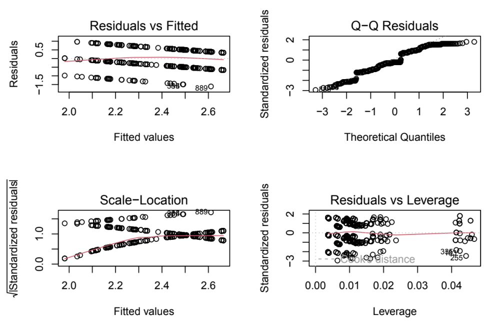


Figure 7: Model 3 Logistic Regression Coefficients

Our fourth and fifth linear regression models looked at how students' final placement into Eng Comp 1, 2, or 3 are impacted by whether or not their case required a 2nd rater. The diagnostic plots for model 5 are much better than model 4, again this is something we found acceptable because nothing in model 4 was significant.



(a) Diagnostic plot for model 4



(b) Diagnostic plot for model 5

Our fourth model looks solely at how sec_grader impacts final placement scores. Model 4 shows that our created variable has a barely positive effect, with a coefficient of 0.03094. However, the p-value of 0.408 shows that the **students' need for a second rater does not significantly impact their final placement.** Figure 9 below supports this, as the confidence interval for the sec_grader variable crosses zero.

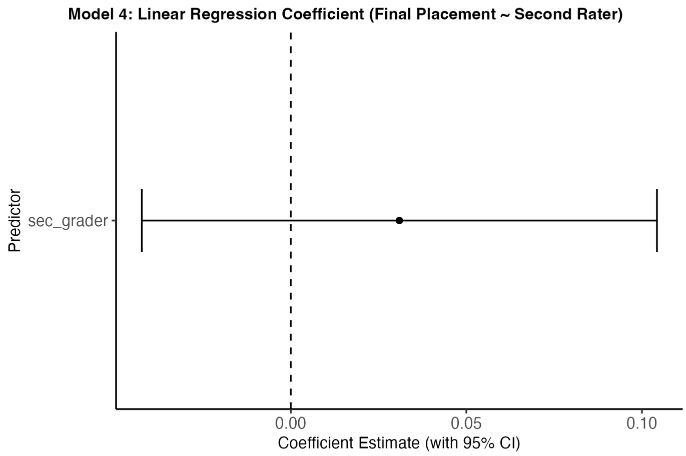


Figure 9: Model 4 Linear Regression Coefficients

Our fifth model considers the same elements as model 4, but with the addition of demographic analysis. This model shows that, at a = 0.01, URM status is a significant variable. The coefficient for this variable, -0.3282, indicates that **students who are underrepresented minorities are more likely to be placed into lower-level Eng Comp courses.** In addition, Pell Grant status is significant at the a = 0.05 level, but not at 0.01. The coefficient of -0.1109 indicates that **students who have received a Pell Grant are also more likely to be placed into lower-level Eng Comp courses.** While this effect is more modest than that of URM status, it still highlights a potential influence of socioeconomic background on placement outcomes.

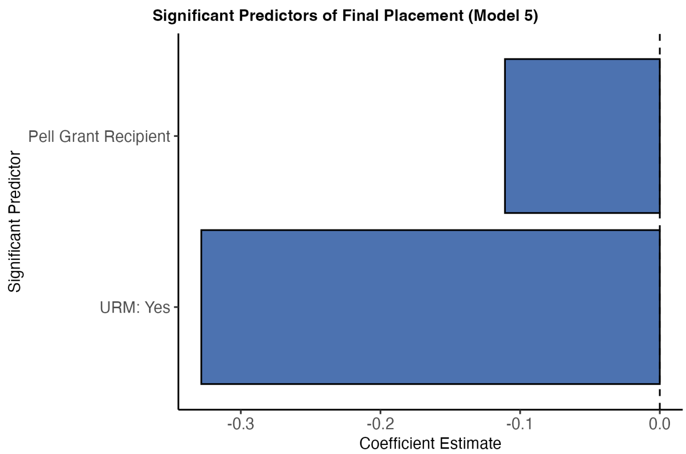


Figure 10: Bar Plot of Significant Predictors for Model 5

Our sixth and final model builds from the conclusions of model 5, considering how demographics impact post IPP survey scores. This model shows three significant predictors at a = 0.01: sec_grader, first generation status (not first gen), and residency status (international). Sec_grader has a negative coefficient, indicating that students who required a second rater had lower survey scores, again aligning with the idea that students with lower self-confidence are more likely to have a mismatch between their course preference and IPP score. Residential status also has a negative coefficient. International students tend to score lower on the post IPP survey, possibly due to language and cultural barriers impacting their confidence. Lastly, first-generation status has a positive coefficient, indicating that students who are not first-generation scored higher, suggesting parental education may boost academic self-confidence. These results will be covered more in the discussion section.

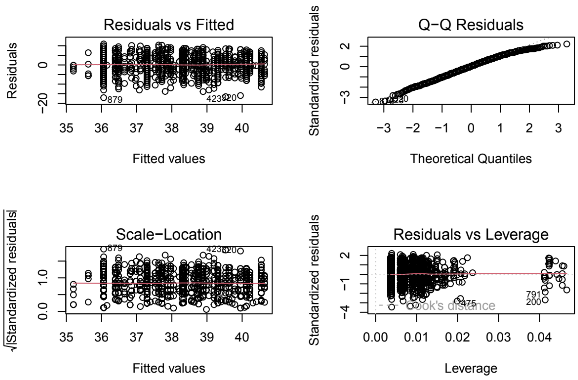


Figure 11: Diagnostic plot for model 6

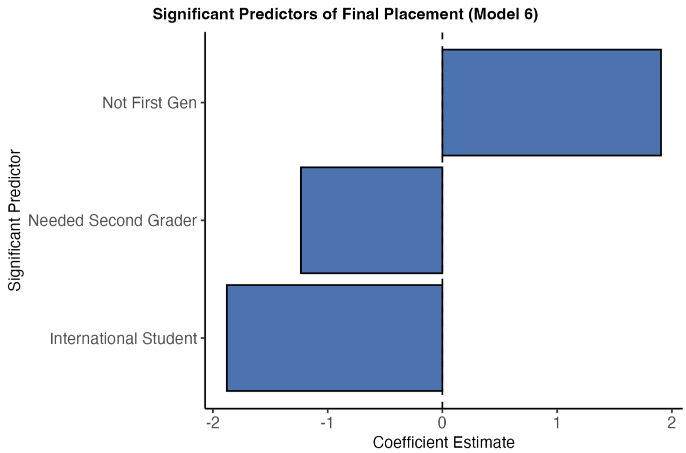


Figure 12: Bar Plot of Significant Predictors for Model 6

6 Discussion

Our first area of interest with this dataset was regarding how student survey scores impacted the need for a second rater. This was meant to be a look at student confidence, considering whether higher survey scores (which correlate to higher student confidence in their writing ability) have an impact on whether the students' IPP score and course preferences matched. More specifically, we were curious if higher levels of student confidence led to more accurate matches between IPP scores and course preferences; if a student with higher confidence would be more likely to accurately discern which Eng Comp level they are at. We also considered the impact of demographics on whether or not the student required a second rater.

Our models show that the odds of a student needing a second rater decrease as post-IPP survey scores increase. This implies that student confidence is a significant factor when considering how students judge their own academic ability. Looking at various demographic factors, we can see that a students' background did not significantly affect their need to have a second rater.

Our second area of interest looks at the impact of needing a second grader on a student's final placement into Eng Comp 1, 2, or 3. In this analysis, we aimed to consider the impacts of mismatches between IPP scores/course preferences and demographics on final placement. We found that a students' need for a second rater did not significantly impact their final score, but their status as an underrepresented minority did. This implies that students who are underrepresented minorities are more likely to be placed lower in their final placement. This observation can be further explained by a review of the literature on this topic.

In our initial analysis of available literature on this subject, we found a multitude of sources discussing the variance in academic self-confidence resulting from various demographic factors, including being underrepresented minorities, first generation, and low income (2). Much of the information in these sources contradicted itself, but we found evidence stating that, despite academic performance not being low among these groups, academic self-confidence often is. As such, our hypothesis for both areas of research was that demographics would impact whether students needed a second rater and their final scores, causing a lack of academic self-confidence and possibly lower final placements due to historical and socio-cultural factors (for example, a lack of resources at schools with less funding, which are often in areas containing underrepresented minorities due to a history of redlining).

However, our analysis proved to show the opposite: demographics did not impact a students' need for a second rater or their final scores as much as we expected. We found that only one predictor (urm) had a significant impact on final scores, and that no demographic predictor had an impact on the need for a second rater. The impact of the urm predictor on final placement is very important to consider, as it calls into question if any accommodations should be made to alleviate any systemic issues.

This discussion called into question how demographics impact student confidence, compelling us to run one more model looking exactly at that. Student confidence, in this case, is measured through the post-IPP survey, which looks at student confidence in their own writing. In our analysis we found that the mismatch between IPP scores and course preferences, as well as a student's international status negatively impacts survey scores. On the other hand, a student having one or more parents/guardians that completed a college degree leads to higher survey scores. These results are highly supported by our exploration into the literature related to this topic.

7 Conclusions and Limitations

This study aimed to determine what factors may affect student placement into English Composition Courses, specifically as they pertain to demographics and student self-confidence. Through the use of linear and logistic regression models, we discovered that demographics had no impact on whether or not a student required a second grader. However, under-represented minorities were found to be more likely to be placed in lower level English courses. These findings contradicted our initial hypothesis, but aligned with previous research that found that under-represented minorities tend to have low academic self-confidence.

While the models were well-suited to our research questions, we must also consider limitations of the experiment. Due to the nature of the IPP exam grading, we must consider the potential for implicit bias in the readers. Since not all exams were assessed by the same reader, there is room for individual bias, especially toward non-english speakers, nonresidents, and under-represented minorities. While readers do not have access to information about students' demographics, they can each student's name, which may lead to assumptions that influence scoring. Such bias could affect the study's findings by making certain demographic factors appear more significant than they truly are. Furthermore, because reader assignment is not a fixed variable, the generalizability of the study is limited.

In the future, research on this topic should aim to reduce bias by keeping grading as controlled and consistent as possible. Furthermore, future studies could follow up with students to compare the grades received with the course they were placed in. Examining student performance could provide insight into the accuracy of student placement as well as potential differences in performance across demographics. For instance, do student's who's exams needed a second grader tend to perform better or worse than those who didn't? Or, do under-represented minorities tend to perform better or worse based on which class they were put in? Looking at student performance after taking the course may also ensure the accuracy of the IPP exam and allow for further analysis of the role of self-confidence in student placement and performance.

References

[1] UCLA's Informed Placement Process (IPP). (n.d.). UCLA Writing Programs. https://www.wp.ucla.edu/placement/ipp/

[2] Penrose, A. M. (2002). Academic Literacy Perceptions and Performance: Comparing First-Generation and Continuing-Generation College Students. Research in the Teaching of English, 36(4), 437–461. http://www.jstor.org/stable/40171585

[3] Owens, A. (2018). Income Segregation between School Districts and Inequality in Students' Achievement. Sociology of Education, 91(1), 1–27. http://www.jstor.org/stable/26428640

[4] Santiago, A. M., & Einarson, M. K. (1998). Background Characteristics as Predictors of Academic Selfconfidence and Academic Self-Efficacy among Graduate Science and Engineering Students. Research in Higher Education, 39(2), 163–198. http://www.jstor.org/stable/40196292

[5] Huang, C. (2013). Gender differences in academic self-efficacy: a meta-analysis. European Journal of Psychology of Education, 28(1), 1–35. http://www.jstor.org/stable/23421703