1-2020

Roadmap to Enrolling Diverse Law School Classes, Volume 4: Contextualizing Admission Factors

AccessLex Institute

Roadmap to Enrolling Diverse Law School Classes

VOLUME 4: Contextualizing Admission Factors

January 2020

AccessLex INSTITUTE®

01 Introduction

Law student diversity is critical to the robust exchange of ideas that is the basis of legal education. Unfortunately, many law schools struggle to enroll classes that reflect the demographics of the regions, states, and even cities in which they are located. A commonly cited reason for the dearth of diversity in many schools is that the pool or "pipeline" of eligible prospective students is not diverse itself.[1]

The premise of this critique is rooted in the manner in which "merit" in the admission process is conceived. LSAT scores are the most prominent admission factor. Past academic performance, typically undergraduate GPA (UGPA), is often given prominence as well. Applicants with high LSAT scores and UGPAs will have many possible options for law school. Law schools also consider other factors, such as personal statements, letters of recommendation, and work experience. But the LSAT score and UGPA typically bear the most weight.[2]

Law schools want to admit and enroll students who are likely to be successful.[3] The most common conception of a successful student is one who does well (or at least adequately) academically, passes the bar exam the first time, and secures full-time, professional employment shortly after graduation, if not before. So, the consideration of applications is essentially an exercise in prediction. Law schools are attempting to identify applicants most likely to attain favorable outcomes.

Ideally, the weight conferred to admission factors is justified by the value of each factor (or combination of factors) in predicting outcomes. Fairness, equity, and the fostering of diverse student bodies require that admission factors be properly contextualized. Narrow conceptions of merit can place applicants from underrepresented backgrounds at a disadvantage in the admission process, due, most significantly, to clear and enduring racial, ethnic, and socioeconomic LSAT score disparities. When factors such as LSAT scores are given undue weight, their impact is unduly harmful to some applicants.

But how can a law school determine the predictive value of different admission factors? The purpose of this fourth volume of the *Roadmap to Enrolling Diverse Law School Classes* series is to provide law schools with a guide to using empirical research methods, such as regression analyses, to gain a better understanding of the relative impact of admission factors on student outcomes. The volume presents the following steps to conducting these analyses:

- Step 1: Consider outcomes to predict.
- Step 2: Choose study subjects.
- Step 3: Contemplate relevant factors.
- Step 4: Collect data.
- Step 5: Calculate relationships between variables.

1 Aaron N. Taylor, 2017, For Diversity: Let's Talk Less About Pipelines and More About Why Blacks Are Not Admitted. http://nationaljurist.com/national-jurist-magazine/diversity-lets-talk-less-about-pipelines-and-more-about-why-blacks-are-not.
2 Vernellia R. Randall, 2006, The Misuses of the LSAT: Discrimination against Blacks and Other Minorities in Law School Admissions. https://heinonline.org/HOL/P?h=hein.journals/stjohn80&i=43.
3 Alexia Brunet Marks and Scott A. Moss, 2015, What Predicts Law Student Success? A Longitudinal Study Correlating Law Student Applicant Data and Law School Outcomes. https://doi.org/10.1111/jels.12114.

02
Conducting an Analysis

A recent report found that 49 percent of the undergraduate institutions surveyed had not conducted predictive validity studies of the factors they considered when reviewing applications.[4] This means that almost half of the schools had little to no insight into the extent to which their admission criteria translated into student outcomes. Without this insight, they are selecting winners and losers in their admission processes without the benefit of critical information. Many law schools are in similar positions.

The following is an illustration of the process of conducting a predictive validity study. The illustration centers on a statistical technique called a *simple regression analysis*. This type of analysis seeks to estimate the value of an outcome variable (e.g., law school GPA) based on the value of a predictor variable (e.g., LSAT score). The examples we use are considered "simple" because each includes only one predictor variable (e.g., LSAT score). In the real world, a rigorous test of predictive validity would use *multivariate regression analysis*, in which the impact of multiple predictor variables would be analyzed in the same regression model. Multivariate regression analysis is essentially a form of predictive modeling, whereby you can estimate both the individual and collective impact of different variables. But for clarity sake, we will rely on a simple regression framework in our illustrations.

The mechanics of conducting this exercise require the involvement of people who are adept at empirical research methods. So, you should retain an expert, if you are not one yourself. But the concepts underlying the exercise can be grasped by most anyone. And it is important that those who would be using the data to inform admission policies understand how to interpret the information the methods yield.

Step 1: Consider Outcomes to Predict

Law schools tend to be most interested in "bottom line" outcomes, like law school grades (LGPA) and bar exam performance. Other, more indirect outcomes, such as student engagement, can also be important. Identifying the outcomes on which to focus will help you devise the research questions that will guide your analyses. In order for the research questions to be testable, the outcome(s) of interest must be a number (or able to be reliably converted to a number—there will be more on this later in Step 3).

For example, if final LGPA is an outcome of interest, you would structure your research question as follows:

> *To what extent does [a certain admission factor] predict final LGPA among [a certain pool of subjects]?*

4 National Association of College Admissions Counseling, 2016, Use of Predictive Validity Studies to Inform Admission Practices. https://www.nacacnet.org/globalassets/documents/publications/research/testvalidity.pdf.

A single study will encompass more than one outcome and, therefore, consist of more than one research question. So, if you happen to also be interested in predictors of first-time bar exam performance, you could include the following research question:

> *To what extent does [a certain admission factor] predict first-time bar exam performance among [a certain pool of subjects]?*

These research questions are incomplete in their current form. To complete them, you will have to identify the pool of study subjects (Step 2) and the admission factors you will study (Step 3).

Step 2: Choose Your Study Subjects

The study subjects are the people whose characteristics, behaviors, experiences, and outcomes will be studied. The outcomes of interest identified in Step 1 will be important in determining who the study subjects should be. If the goal is to predict final LGPA, then the pool of study subjects must consist of people who have completed law school. If the goal is to predict bar exam performance, then the pool must consist of people who have taken the bar exam. If both final LGPA and bar exam performance are outcomes of interest, then a pool of graduates of your school who have taken the bar exam would be the ideal study subjects.

There will be remaining questions, such as whether you should study just one year of graduates or multiple years. Typically, it is better to study multiple (say, three) years of graduates. Then, there are questions regarding whether you study each year individually or combine the data and study all subjects as one group. Do not fret much over these considerations at this stage. The researcher you get to conduct the study will advise you regarding the wisdom of different approaches.

For purposes of this illustration, let us say you are studying the last three years of graduates as one group. With that, the sample research questions continue to develop:

> *To what extent did [a certain admission factor] predict final LGPA among 2015, 2016, and 2017 graduates of the law school?*
>
> *To what extent did [a certain admission factor] predict first-time bar exam performance among 2015, 2016, and 2017 graduates of the law school?*

Step 3: Contemplate Relevant Factors

This step involves a consideration of factors that possibly help predict the outcomes you identified in Step 1. You do not have to try and identify every factor that might have predictive value. In fact, it is impossible to do so. But reasonable effort should be made to identify and study all factors that play tangible roles in the consideration of applicants. Other potentially relevant factors should also be considered.

The LSAT score and UGPA are important and obvious factors to investigate. There are other possible factors; some may not be numerical in their raw form. For example, if a school wanted to investigate whether personal statements or letters of recommendation had predictive value, it might develop rubrics that could then be used to assign numerical scores to statements and letters. Doing so would require additional effort and coordination. But the exercise could very well be worth it in ensuring that the process is rooted in multifaceted evidence of what matters.

The number of factors identified as potentially relevant will influence the number of research questions. There will be fully formed research questions for each factor. For example, if LSAT and UGPA are two factors, the following could be the research questions:

To what extent did LSAT scores predict final LGPA among 2015, 2016, and 2017 graduates of the law school?

To what extent did LSAT scores predict first-time bar exam performance among 2015, 2016, and 2017 graduates of the law school?

To what extent did UGPAs predict final LGPA among 2015, 2016, and 2017 graduates of the law school?

To what extent did UGPAs predict first-time bar exam performance among 2015, 2016, and 2017 graduates of the law school?

Step 4: Collect Data

Once potential factors are identified, the next step is to collect and log the numerical outputs (e.g., scores) associated with each factor, for each study subject. For this illustration, that means collecting LSAT score, UGPA, final LGPA, and first-time bar exam result for each member of the 2015, 2016, and 2017 entering classes.

In some instances, data collection is relatively straightforward. LSAT scores and GPA data are already systematically collected by law schools. Collecting other types of data, such as the scored personal statements and recommendations, or even bar exam results, would require more effort and coordination. All data should be compiled in a form that allows for analyses and tabulation, using uniform means of formatting and coding. Data should also be reviewed for errors (e.g., an LSAT score of 190) and "cleaned" where necessary. Missing data (e.g., a study subject without a bar result) may be an issue as well. Your researcher will advise you on how to best handle missing data issues.

Step 5: Calculate Relationships Between Variables

Notice that at this stage, we are referring to factors as "variables." Once your data is collected, coded and cleaned, you will want to explore relationships between variables. As mentioned earlier, your research questions will guide your analyses. In this illustration, you would seek to measure the extent to which the predictor (or independent) variables, like LSAT scores and UGPAs, predict the outcome (or dependent) variables, like final LGPA and first-time bar exam performance. Your researcher will explore these relationships using regression analyses. While you do not have to be adept at running regressions, it is important to understand what the numerical outputs of a regression analysis tell you. In the next section, we will discuss the types of information produced by this method of analysis and how it is relevant to how you contextualize admission factors.

03
Interpreting Results of Analyses

Your regression analyses will yield a range of coefficients and values. These numbers will tell you whether there are "statistically significant" relationships (more on this later) between variables and provide you with insights regarding the nature of those relationships. This section will discuss two types of regression analyses: linear and logit.

Linear regressions are appropriate when the values of both the predictor variables and the outcome variables are numerical and arrayed on a scale. Examples of such variables include LSAT scores, grade point averages, and personal statement quality scores. Logit regressions are appropriate when one or more of the variables yield values that are discrete, meaning they are not arrayed on a scale. They are commonly non-numerical in their raw form. Examples of such variables include bar exam result, which in its raw form is expressed as either Pass or Fail.

Linear Regression Analysis

Below is a sample output from a linear regression analysis measuring the predictive influence of LSAT scores and UGPAs on final LGPA. The data is taken from an actual analysis we conducted, though some of the values have been changed to ensure confidentiality. Not all outputs will look exactly like this, but they will consist of much of the same information.

Table 1: Linear Regression Results for Final LGPA

Variable	Coefficient (P-Value)	Coefficient (P-Value)
LSAT Score	0.028* (<0.001)	0.026* (<0.001)
UGPA		0.0 (<0.001)
Intercept	-1.410 (0.092)	-2.127* (0.011)
Number of Observations	152	148
R^2	0.124	0.203
Adjusted R^2	0.119	0.194

* indicates statistical significance

The output consists of a lot of seemingly random information. But most of it is tangibly useful. Let's walk through it.

Variable	**Coefficient (P-Value)**	**Coefficient (P-Value)**
LSAT Score	0.028* (<0.001)	0.026* (<0.001)

The first column identifies the LSAT score as the variable being analyzed. The middle column provides the regression coefficient, which is 0.028. Before we get to the meaning of the coefficient, let's talk about both the asterisk displayed next to the coefficient and the "P-Value" (<0.001) below it. The asterisk tells us that the regression coefficient is statistically significant, which means that the observed relationship between the variable and the outcome (as represented by the coefficient) is very likely an actual relationship and not a random association.

So, in this case, the asterisk tells us that LSAT score very likely correlated with final LGPA. The p-value estimates the actual odds that the relationship is random. In this case, the p-value of <0.001 means that there is less than a one-tenth of one percent chance that the relationship is random. This represents a high level of statistical significance. A lack of statistical significance does not necessarily mean that the observed relationship lacks practical significance; it means that the regression model was unable to rule out randomity with sufficient statistical certainty.

The coefficient in this table provides a directly interpretable estimate of the effect of a one-unit change in the predictor variable on the outcome variable. In this case, the coefficient in the middle column (0.028) means that a one-unit (or one point) change in LSAT score resulted in a 0.028 change in final LGPA among the study subjects. The fact that the coefficient is a positive number means that the unit change and the impact flow in the same direction. So, an increase in LSAT score was associated with an increase in final LGPA. Conversely, a decrease in LSAT score resulted in a decrease in final LGPA, Had the coefficient been negative (e.g., -0.028), the unit change and impact would flow in opposite directions. A one-unit increase in LSAT score would have resulted in a decrease in final LGPA and vice versa.

Let's ignore the far-right column for the LSAT variable for now and take look at the output for the UGPA variable:

Variable	**Coefficient (P-Value)**	**Coefficient (P-Value)**
LSAT Score	0.028* (<0.001)	0.026* (<0.001)
UGPA		0.217* (<0.001)

The asterisk and p-value tell us that the coefficient (0.217) is statistically significant, with less than a one-tenth of one percent chance that the relationship is random. The coefficient tells us that a one-unit change in UGPA resulted in a 0.217 change in final LGPA. This is where the ability to provide context to the data is very important. A 0.217 coefficient is very large in this context. To a person with knowledge of law admission, it would likely look immediately suspect that UGPA would have such impact. And it should. While the coefficient is accurate, it is misleading. The reason for this is the way the model defines a one-unit change in UGPA. The model defines a one-unit change as a whole grade point (e.g., 2.0 to 3.0). A whole-point change may make sense for the LSAT variable, given that scores are expressed in whole points and are on a 60-point scale, but such framing is not useful for GPAs.

A more useful conception of a one-unit change in GPA would be a one-tenth change (e.g., 3.0 to 3.1), which would be associated with a coefficient of 0.0217 (one-tenth the size of the original estimate). So, a one-tenth change in UGPA resulted in a 0.0217 change in LGPA. Your researcher will be an expert in running these analyses but may not have contextual knowledge of law admission or legal education. Therefore, it is important that you are able to interpret the data your researcher provides and add context, where appropriate.

Let's now go back to the far-right column for the LSAT score variable and include the UGPA variable:

Variable	**Coefficient (P-Value)**	**Coefficient (P-Value)**
LSAT Score	0.028* (<0.001)	0.026* (<0.001)
UGPA		0.127* (<0.001)

The data in the far-right column capture the influence and interplay of multiple predictor variables on the outcome variable. In this model, the influence of both the LSAT score and UGPA on final LGPA were analyzed. This is a classic multivariate regression analysis. Looking at the LSAT score row, you should notice that the LSAT coefficient falls slightly (from 0.028 to 0.026) when UGPA is added to the analysis (far-right column). This is due to there being aspects of the LSAT score's solitary relationship with LGPA that were usurped by the UGPA once it was added. You can conceive of a multivariate regression analysis as a set of individual variables competing against each other for "credit." As relevant variables are added to the model, coefficients for the other variables fall. So, if personal statement scores were added to this analysis, and they had a relationship with LGPA, the coefficients for LSAT score and UGPA would probably fall by at least a little.

So, let's talk about the rest of the output.

Table 1: Linear Regression Results for Final LGPA

Variable	**Coefficient (P-Value)**	**Coefficient (P-Value)**
LSAT Score	0.028* (<0.001)	0.026* (<0.001)
UGPA		0.227* (<0.001)
Intercept	-1.410 (0.092)	-2.127* (0.011)
Number of Observations	152	148
R^2	0.124	0.203
Adjusted R^2	0.119	0.194

* indicates statistical significance

Do not worry about the Intercept. It does not have much practical value in this context. The row labeled "Number of Observations" is the number of study subjects whose data was analyzed. The R^2 and Adjusted R^2 coefficients both tell you the amount of variation in the values of outcome variables that is explained by variation in the values of predictor variables. That is a mouthful, so here are two illustrations:

- The R^2 of 0.124 in the middle column tells you that about 12.4% of the variation in final LGPAs among study subjects was explained by variation in their LSAT scores. That means that LSAT scores had a tangible, but limited influence on LGPA. Almost 90% of final LGPA variation was unexplained by the variation in LSAT scores.
- The Adjusted R^2 of 0.194 in the far-right column tells you that about 19.4% of the variation in LGPAs among study subjects was explained by variation in their LSAT scores *and* UGPAs. Including the UGPAs in the model increased its predictive value, but there was still more than 80% of the LGPA variation that was unexplained by these two factors.

Generally speaking, the R^2 and Adjusted R^2 coefficients measure how well the model performs in estimating (i.e., predicting) the outcome variable. The higher the percentage, the better the performance.

For simple (one factor) regression analyses, you should rely on the R^2 value. But for multivariate analyses, the Adjusted R^2 is better, largely because it better screens statistical "noise." Do not worry if you do not fully understand that. Your researcher will know which value is better for different types of analyses.

Logit Regression

Below is a sample logit regression output, tying various factors to bar exam result (pass/fail). The data was taken from an actual analysis (with some of the values changed).

Table 2: Linear Regression Results for Bar Exam Result

Variable	Coefficient (P-Value)
LSAT Score	-0.009 (0.905)
Undergraduate GPA	-0.6 (0.526)
Final LGPA	8.043* (0.011)
Credit Hours: Bar-tested courses	0.178* (0.027)
Credit Hours: Skill-based courses	0.083 (0.448)
Observations Log Likelihood	148 -28.772
Akaike Inf. Crit.	85.545

* indicates statistical significance

Unlike the linear regression output, the data in the table are not directly interpretable because the coefficients do not represent a one-to-one relationship between the predictor variables and the outcome. As a result, most of the numbers in the logit regression output have little practical value. But there are some aspects of the table that are important and easy to interpret. Like in the linear regression output, the asterisks signify that the associated variables are statistically significant predictors of bar exam passage or failure.

So, in the table, final LGPA and credit-hours completed in bar-tested courses were both statistically significant predictors of whether someone passed or failed the bar exam. The p-values (in the parentheses) again estimate the actual odds that the relationship between each variable and bar exam result is random and *not* significant. For final LGPA, there is a one-tenth of one percent chance that the relationship is random; on the other hand, for LSAT score, there is a 90.5% chance (hence the reason it does not have an asterisk).

The coefficients do not have much practical value on their own, but their directionality (whether they are positive or negative) matters. Because both final LGPA and credit-hours completed in bar-tested courses are positive, this means that the higher the LGPA and the more credit-hours in bar-tested courses, the higher the chance of passing the bar exam. Had the coefficients been negative, the relationships would flow in opposite directions.

Your researcher can increase the tangible usefulness of a logit regression analysis by using the data to calculate predicted probabilities, which show how a change in a predictor variable (e,g., final LGPA) impacts the odds of a particular outcome (e.g., bar exam passage). The true value of these probabilities is in how they account for other variables that are included in the model. For example, a predicted probability analysis may tell you that a student with a final LGPA of 2.50 has a 40% chance of passing the bar exam, while a student with a 3.00 LGPA has a 70% chance. And in estimating these probabilities, the model is holding constant other factors, such as the number of credit-hours taken in bar-tested courses. The goal is to isolate the impact of individual factors on the outcome.

04
Using Data to Inform Admission Consideration

A primary purpose of conducting these analyses should be to use the resulting data to help inform how applicants for admission are considered. What factors should be considered and to what extent should they impact admission decisions are important questions. The manner in which most law schools consider applications is not overly formulaic; this is largely because most factors are not quantitative or numerical in nature. Factors such as personal statements and letters of recommendation comprise of words. Reviewers typically read these items and make subjective judgments regarding their quality. A similar process occurs when considering other factors, such as applicant background and demographics.

The relative lack of ostensibly objective admission factors contributes to overreliance on the LSAT score when considering applicants. The LSAT is a uniform test that is scored on a uniform scale. It is arguably the most objective admission factor. The UGPA also enjoys favored, and often inflated, status in the admission process. Applicant UGPAs are all converted by LSAC to a uniform scale, adding an element of objectivity, even though the underlying grades are earned at different schools from different professors with different grading practices and across different academic majors.

LSAT scores and UGPAs have value as predictors of law school academic performance. Therefore, they are relevant to admission consideration. Their value, however, is limited. Additionally, these factors will likely have no statistically significant relationship to bar exam performance at most schools (though indirect relationships exist). This means that most of what determines how a student performs in law school and on the bar exam is determined by other factors.

A well-designed process for considering applicants is based on the best information of what factors help predict student success and the extent to which they help predict it. Factors that have only limited predictive value should be given appropriately limited sway in determining who gets admitted and who gets denied. If the LSAT score predicts, say, 12% of the variance in final LGPA (as was the case for the school used in the earlier illustration), it should be treated as just one admission factor among others. But more than that, the collective weight of other factors should outweigh the individual weight of the LSAT. Same for the UGPA. This is often not the case in law school admission.

Schools should strive to measure the predictive impact of all factors that are considered when reviewing applicants. Personal statements and letters of recommendation are typically viewed as inherently subjective, largely because they are difficult to compare across applicants. But the predictive value of these factors on student outcomes can be estimated as well. As mentioned earlier, schools could develop scoring rubrics for these documents, and then tie scores to student outcomes using regression analyses.

Schools could also investigate other factors that may not be systematically considered but are tracked. Examples might include applicants' academic majors, years of work experience, and the possession of graduate degrees. The goal would be to use the information to properly contextualize admission factors and reduce the extent to which gaps in seeminly objective applicant information are filled by an overemphasis of certain factors. How merit is conceived in the admission process should be as multifaceted as our applicants.

AccessLex Institute®, in partnership with its nearly 200 nonprofit and state-affiliated American Bar Association-approved Member law schools, has been committed to improving access to legal education and to maximizing the affordability and value of a law degree since 1983. The AccessLex Center for Legal Education Excellence® advocates for policies that make legal education work better for students and society alike, and conducts research on the most critical issues facing legal education today. The AccessLex Center for Education and Financial Capability® offers on-campus and online financial education programming and resources to help students confidently manage their finances on their way to achieving personal and professional success. AccessLex Institute is a nonprofit organization with offices in West Chester, PA and Washington, D.C., and with accredited financial counselors throughout the U.S.

Corporate Office:	10 N. High Street, Suite 400 West Chester, PA 19380
Washington, D.C. Office:	440 First Street NW, Suite 550 Washington, D.C. 20001

AccessLex.org

AccessLex INSTITUTE®

EMPOWERING THE NEXT GENERATION OF LAWYERS®