A Method for Choosing Weights to Predict College Grades for Admission Decisions and to Assess their Fairness by Race/Ethnicity

Steven Andrew Culpepper University of Colorado Denver
Ernest C. Davenport University of Minnesota
Mark L. Davison University of Minnesota

Previous research suggests that equal weights tend to outperform statistically optimal weights in cross-validation studies. This paper argues that the findings from the equal weights literature are relevant for researchers that predict college grades and/or assess differential prediction of college grades by student characteristics. An application of the criterion profile methodology (CPM) is presented to demonstrate how to examine individual criterion profiles. This study showed how to use the CPM to determine the extent to which equal and statistically optimal coefficients differentially predicted college grades for minority and majority students. The results support previous findings, in that, 92.5% of the explained variance in college grades was attributed to equal weights, where standardized test scores and high school rank were weighted equally, and 7.5% of the explained variance was accounted for by statistically optimal coefficients that weighted ACT Math scores less than ACT English and high school rank. Additionally, equally weighting admission information was more accurate for predicting Asian Americans' future academic performance than European Americans.

Prediction is an important aspect of scientific endeavors. For instance, educators predict student achievement, psychologists classify clients into different diagnoses, university personnel place students into developmentally appropriate courses, and economists forecast stock prices and economic conditions. Multiple regression is an important statistical tool for making these types of predictions.

One classic application of multiple regression is the prediction of college grades with standardized test scores and measures of high school academic success (Hills, 1964; McKelpin, 1965; Munday, 1965; Richards & Lutz, 1968; Sassenrath & Pugh, 1965; Stanley & Porter, 1967). Many applications of multiple regression for predicting college grades are designed to create an equation (which is later referred to as a selection equation) for admission officers to objectively select or sort applicants based upon predicted academic performance. However, practitioners interested in prediction should be cautious about using regression equations to predict applicants' grades. Specifically, using an estimated regression equation to select or sort applicants assumes the intercept and slope coefficients are statistically equal (i.e., invariant) across different demographic groups (Dorans, 2004).

Differential prediction, which occurs when the intercept and/or slope coefficients differ between two or more groups (Cleary, 1968; Humphreys, 1952), can impact admissions decisions that use quantitative information and multiple regression predictions. For instance, intercept differences indicate that one group's academic performance is over-predicted and another group's performance is under-predicted. Previous research suggests that racial/ethnic minority group performance is consistently over-predicted with a common regression equation (Breland, 1979; Burton & Ramist, 2001; Cleary, Humphreys, Kendrick, & Wesman, 1975; Duran, 1983; Linn, 1973; Stanley & Porter, 1967; Wilson, 1983; Young, 2001). In this case, admission decisions are biased in favor of minorities; using one equation to predict academic performance will tend to over-predict minorities' actual performance. In contrast, female's performance is consistently under-predicted (Bridgeman, McCamley-Jenkins, & Ervin, 2000; Chou & Huberty, 1990; Elliott & Strenta, 1988; Noble, 2003; Pennock-Román, 1994; Ramist, Lewis, & McCamley-Jenkins, 1993; Young, 1994).

Slope differences provide evidence about the accuracy of different variables for predicting group outcomes. For example, Young's (2001) review of differential prediction studies suggests that minority students tend to have smaller slope coefficients for academic predictors, such as test scores and high school grades, than their European American counterparts. Slope differences by race/ethnicity are often evidence that test scores and/or high school grades are less valuable indicators of minority students' future academic performance and should be considered differently and perhaps with less value in admissions decisions.

Researchers and practitioners need to assess the presence of differential prediction to ensure responsible multiple regression-based predictions. While multiple regression is used to assess differential prediction (the application of regression to assessing differential prediction is also referred to as moderated multiple regression, or MMR, see for example Saunders, 1956), research suggests that it is not a perfect tool for uncovering differential prediction. Aguinis and Stone-Romero (1997) note that the power to detect slope coefficient differences is smaller when certain artifacts are present (e.g., small sample sizes, relatively low representation of minorities in the sample, measurement error in the predictor and criterion, and range restriction in the predictor). That is, the likelihood of uncovering true differences in slope coefficients is less likely when the aforementioned artifacts are present. This translates into researchers being less likely to conclude that test scores and/or high school academic achievement are less valuable indicators for a given minority group (e.g., comparisons by race/ethnicity, gender, first-generation status, etc.) even when such an inference is true in the population.

Dana and Dawes (2004) offer an additional caution for statisticians and practitioners who make predictions with multiple regression. Through the use of statistical simulations, Dana and Dawes (2004) contributed evidence to an existing body of research (Davis & Sauser, 1991; Dawes, 1979; Dawes & Corrigan, 1974; Dawes, Faust, & Meehl, 1989) that equal weights (i.e., slope coefficients that are equal for every predictor) tend to outperform statistically optimal weights (i.e., standardized slope coefficients derived with ordinary least squares, or OLS) in cross-validation samples. Dana and Dawes (2004) boldly conclude that multiple regression should not be used for prediction purposes when the total variance explained is small (i.e., R^2 is less than 0.25). Similarity, Einhorn and Hogarth (1975) suggest that regression should not be used when $R^2 < 0.50$.

High school grades and standardized test scores tend to account for less than 50% of the total variation in college grades (and sometimes less than 25% of variation). This poses a challenge for researchers who wish to predict college grades and/or conduct differential prediction studies. In particular, differential prediction studies use MMR to test the extent to which subgroups' differ in statistically optimal slope coefficients. Dana and Dawes' (2004) findings provide evidence that comparing the equivalence of statistically optimal weights may be inadvisable when R2 is small. Indeed, if equal weights account for the majority of variance in college grades rather than statistically optimal weights, it may be more appropriate to determine the extent to which equal weights differentially predict college grades for different subgroups.

The goal of this paper is to show how an external profile analysis technique can be used to assess differential prediction by race, while concurrently examining the value of equal and statistically optimal weights. A simple application of a criterion profile methodology (CPM; Davison & Davenport, 2002) is presented to test the extent to which statistically optimal and equal weights differentially predict the college grades of two minority groups (African; Asian Americans) when compared to European Americans.

One goal of this study was to assess differential prediction in a way that accommodates Dana and Dawes' (2004) concerns. Accordingly, the first section of this paper introduces the mathematical formulation of the CPM and describes how the CPM can address Dana and Dawes' (2004) concerns.

The CPM is an external profile analysis technique (Davison & Davenport, 2002) that yields a predictor profile that differentiates between subjects with high and low scores on a criterion, such as college grades. So, another goal of this study was to introduce researchers to the CPM, since it is applicable to other research endeavors. The profile analysis feature of the CPM was described to clearly articulate the model. In this study, high school grades and ACT test scores were used to demonstrate individual and sub-group profile differences. The third section discusses the results of the differential prediction analysis as it relates to using the CPM and the last section provides concluding remarks.

Description of the Criterion Profile Methodology

Profiles patterns identified with cluster analysis or multidimensional scaling have been criticized for not exhibiting criterion-related validity evidence (Watkins, 2000). The advantage of the CPM is that the identified profiles are explicitly related to a criterion, such as college grades in this study, and exhibit some degree of validity as determined by the strength of the relationship between the predictors and the criterion (Davison & Davenport, 2002). That is, the identified predictor patterns distinguish those low and high on the criterion variable.

The CPM parses the variation in a criterion variable explained by a set of independent variables into two components: a level effect, which is characterized by an equally weighted linear composite of the predictors, and a pattern effect, which is the covariance between a subject's predictor profile and the regression coefficients. The relationship between the level effect and equal weights versus the pattern effect and optimal weights is discussed below.

To formally define these two effects we start with the usual regression model in Equation 1 below:

$$Y_p = \sum_{v=1}^{V} \beta_v X_{pv} + a + e_p \tag{1}$$

where Y_p represents the criterion score for person p, β_v represents the regression coefficient for variable v (v = 1 to V where V is the number of predictors and $V \geq 2$), and X_{pv} is the score for subject p on predictor v. Finally, a is the intercept of the regression equation and e_p is a random error term. The criterion profile is defined as the set of slope coefficients, β_1, β_2, … , and β_v, for the p predictors.

Davison and Davenport (2002) prove that the regression model in equation 1 is equivalent to the following model:

$$Y_p = \gamma_1 Cov_p + \gamma_2 \overline{X}_p + a + e_p \tag{2}$$

where, $\overline{X}_p$ is referred to as level, Cov_p is referred to as pattern, and γ_1 and γ_2 are their respective slope coefficients (these coefficients are standardized if Cov_p and $\overline{X}_p$ represent z-scores). The first term on the right of Equation 2 constitutes the pattern effect, and the second term is the level effect. Regardless of the original number of predictors, the original regression equation can be reduced to three terms, a, Cov_p, and $\overline{X}_p$. It is important to note that Y_p and e_p are the same in Equations 1 and 2, so that the level and pattern variables together account for the same proportion of variation in the criterion as the original variables.

The equations for level and pattern are presented below in Equations 3 and 4:

$$\overline{X}_p = \frac{\sum_{v=1}^{V} X_{pv}}{V} \quad ; \tag{3}$$

$$Cov_p = \frac{\sum_{v=1}^{V} (\beta_v - \overline{\beta})(X_{pv} - \overline{X}_p)}{V} \quad ; \tag{4}$$

where the only new variable, $\overline{\beta}$, in Equation 4 represents the average of the regression coefficients. Equation 3 shows that level, $\overline{X}_p$, is person p's unweighted average on the independent variables, which suggests that individuals who tend to have larger (smaller) standardized values on the V predictors will also tend to have larger (smaller) values for level.

Pattern (Cov_p) is the covariance between person p's predictor scores and the weights from the original regression, and it is a measure of the match between the observed score profile of person p and the pattern that distinguishes people with high scores on the criterion. Individuals with predictor scores whose pattern matches the configuration of the regression weights will have larger profile match statistics Cov_p and therefore higher predicted values.

Because it is a covariance measure, pattern is positive for subjects whose scores are consistent with the criterion profile (i.e., the configuration of the slope coefficients, β_v) and negative for subjects whose scores are consistent with the mirror image of the criterion profile. The mirror image profile is defined by slope coefficients with the exact opposite configuration as the criterion profile. The coefficients for the mirror image profile (ψ_v) can be found with the following expression: $\psi_v = \overline{\beta} - (\beta_v - \overline{\beta}) = 2\overline{\beta} - \beta_v$. Subjects with predictor profiles corresponding to the mirror image pattern tend to have lower predicted values after controlling for level. The criterion and mirror image profiles are discussed later during the application of the CPM with admissions data.

Value of the CPM for Differential Prediction Studies

The CPM is useful for identifying profiles that differentiate between individuals with high and low criterion values, such as first year cumulative grade point average in our case. However, the criterion profile is only useful for differentiating between high and low CGPA scorers when the pattern effect is statistically significant after controlling for level, which occurs when the statistically optimal weights add predictive value in addition to equal weights. In fact, pattern is generally statistically significant when variability exists among the standardized regression weights. In cases where pattern provides no additional prediction in the criterion over that of level, the equal weight profile differentiates between low and high scoring subjects, subjects with larger criterion scores tend to have high values on all the predictors rather than a configuration of predictor scores.

The pattern variable is also important, since it represents the extent to which practitioners should employ statistically optimal weights in decisions. Referring to Equation 4 again, we see that pattern is particularly important for assessing differential prediction. The relative size of the regression weights vary along two extremes: the weights are either close to being equal or they differ substantially in magnitude. Pattern will account for less variation in a criterion when the regression weights are similar or equal to each other, which would lead to equal weights outperforming statistically optimal weights. Conversely, statistically optimal weights are important for prediction purposes when pattern accounts for relatively more variation in a criterion than level, or equal weights. Pattern is useful to the extent that the optimal weights (betas) vary and this variance in weights accounts for differences in the criterion.

This study assesses the value of using equal vs. statistically optimal weights by estimating the amount of variation that is accounted for by level and pattern. More formally, the hypothesis is:

$$H_0: R^2_{\text{Level}} = R^2_{\text{Level + Pattern}} \text{ and } H_1: R^2_{\text{Level}} \neq R^2_{\text{Level + Pattern}};$$

which is testable with the traditional F-test comparing parametric regression models with V-1 and N-V-1 degrees of freedom where V is the number of independent variables and N is the sample size (Davison & Davenport, 2002). Substantively, this test will provide evidence for whether or not statistically optimal weights provide predictive value above and beyond equal weights.

Methods

Sample

This study used data from the entering class of 2000 at a public research university. The data was collected from each student during the pre-college admissions process and provided to the researchers by the Office for Institutional Research. The sample consisted of 2,035 students who enrolled in the College of Liberal Arts (CLA) fall 2000 and persisted through one year of academic study. These 2,035 students were disaggregated by self-reported race/ethnicity. Of the 2,035 students, 68 were African American (AFA), 11 were American Indian, 186 were Asian American (ASA), 1,683 were European American (EA), 38 were Hispanic, 5 were International, and 44 were unidentified. Only the AFA, ASA and EA groups were included in the analyses, since there were small numbers of American Indians, Hispanics, International, and unidentified students in the sample. The final sample size included 1,933 students (four students had missing scores on at least one of the predictors).

Variables

In this application of the CPM, the predictors of interest were students' ACT English sub-score (ACTE), ACT math test sub-score (ACTM), and high school percentile rank (HSR) and the criterion was first-year cumulative grade point average (CGPA). The regression and CPM analyses were conducted by standardizing the predictors and criterion across racial subgroups onto a z-score scale with a mean of zero and variance of one. Table 1 reports descriptive statistics of the variables. Certainly, one could argue that ACTE, ACTM, and HSR predict first-year grades differently depending upon the type of coursework or degree program in which a student engages. It is important to note that the students in the CLA were chosen to reduce the potential heterogeneity in regression equations across different colleges within the university.

Point of Caution

With respect to the CPM, there are no limitations of formulating the general linear model in terms of level and pattern. In fact, the reconfiguration of the general linear model into the CPM accounts for the same proportion of variance in the criterion. However, it is important to consider

Table 1. Multiple Regression and Criterion Profile Methodology Summary.

	Std. Weight	Sig.	Part Corr.	*M*	*SD*
Regression Model[a]					
ACT English	0.306	***	0.261	24.3	4.30
ACT Math	0.087	***	0.074	24.6	4.08
High School Rank	0.306	***	0.302	80.8	11.50
CPM Model[a]					
Level	0.496	***	0.493	-0.0004	0.71
Pattern	0.138	***	0.137	-0.00003	0.05
Cross-Validation Summary	Sample 1		Sample 2		
R^2, Level Only	0.233		0.229		
R^2, Level + Pattern	0.259		0.242		

Note: Std Weight = Standardized slope coefficient, Sig. = Level of significance,
Part Corr. = Part Correlation, *M* = average, *SD* = standard deviation.
The average CGPA was 3.01 with a standard deviation equal to 0.59.
[a]The model R^2 = 0.250 for both the regression and CPM models.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

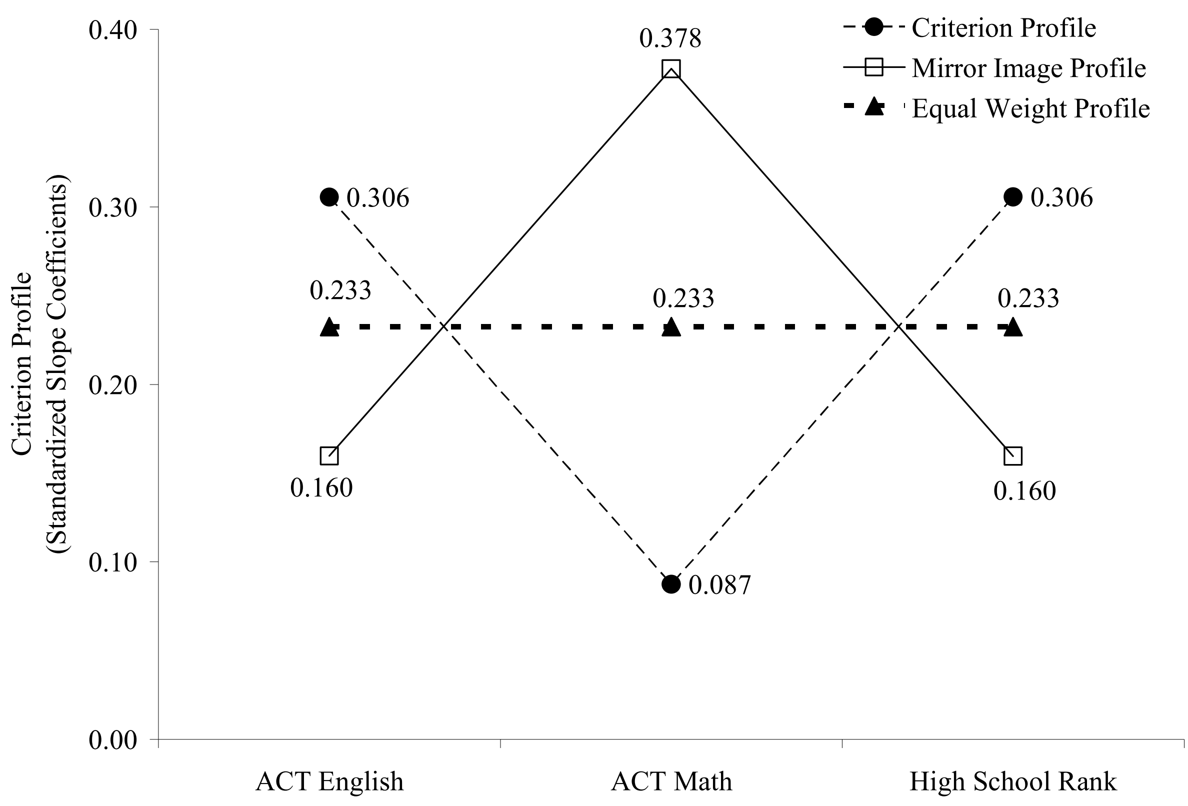

Figure 1. Pooled, within race/ethnicity, and equal weight criterion profiles.
Note. R^2 = 0.250 for the criterion profile and R^2 = 0.231 for the equal weight profile.

one issue to ensure meaningful CPM analyses, which is that the independent variables need to be on either a substantively meaningful scale, such as the number of credit hours in various mathematics courses (Davison & Davenport, 2002), or the same scale, such as z-scores, to yield regression coefficients that are comparable in the criterion profile. Failure to address the scaling of the predictors may produce misleading or substantively uninformative results (Davison & Davenport, 2002).

Results

The results section consists of two subsections. In an effort to further articulate and demonstrate the CPM, the first section applies the CPM to identify a profile pattern that distinguishes those high and low on the criterion variable. The second section uses the CPM to assess the extent to which equal and statistically optimal weights differentially predict CGPA.

Profile Application of the Criterion Profile Methodology: Understanding Individual Differences.

Table 1 presents the regression summary of ACTE, ACTM, and HSR as predictors of CGPA. The regression model accounted for approximately 25.0% of the variation in CGPA. Furthermore, all the predictors were positively related to CGPA and were statistically significant at the 0.001 level. The standardized slope coefficients equaled 0.306, 0.087, and 0.306, for ACTE, ACTM, and HSR, respectively. The standardized slope coefficients define the criterion profile, which was characterized by larger weights for ACTE and HSR than for ACTM. The mirror image profile consisted of weights with an exact opposite configuration of the criterion profile. Figure 1 plots the criterion and mirror image profiles, in addition to the equal weight profile.

Figure 2 plots three subjects' standardized predictor profiles to demonstrate how the CPM can be used to describe individual differences. Figure 2 shows that subject 261 more closely matched the criterion profile and subject 144 matched the mirror image profile. The average pattern was approximately zero with a corresponding standard deviation of 0.05. Subject 144 had a pattern value about two standard deviations below the mean (Cov_{144} = -0.13) and subject 261 had a pattern value about two standard deviations above the mean (Cov_{261} = 0.12)**.** Subject 122's standardized predictor profile resembled the equal weight profile, since the profile was nearly flat and the values for the three predictors were within one standard deviation of each other. Additionally, subject 122's profile did not match either the criterion or mirror image profiles, as indicated by Cov_{122} = -0.001. Moreover, subject 122 had the largest level value and subject 261 had the smallest value for level.

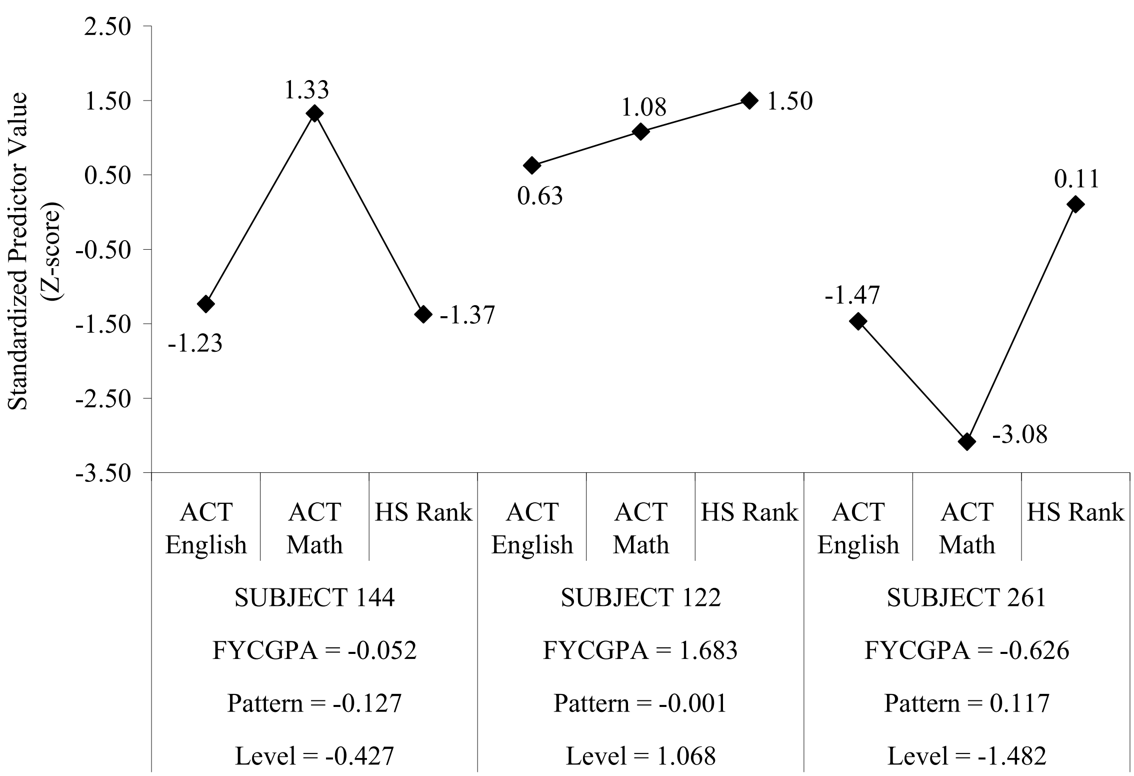

Figure 2. Three subjects' predictor profiles.

Figure 2 also includes information pertaining to the subjects' CGPA, which were standardized to z-scores. Of these three subjects, 122 performed the best academically (1.68) followed by Subject 144 (-0.05), and Subject 261 (-0.63). It is important to determine the extent to which higher academic performance was associated with individual differences in level, individual differences in pattern, or

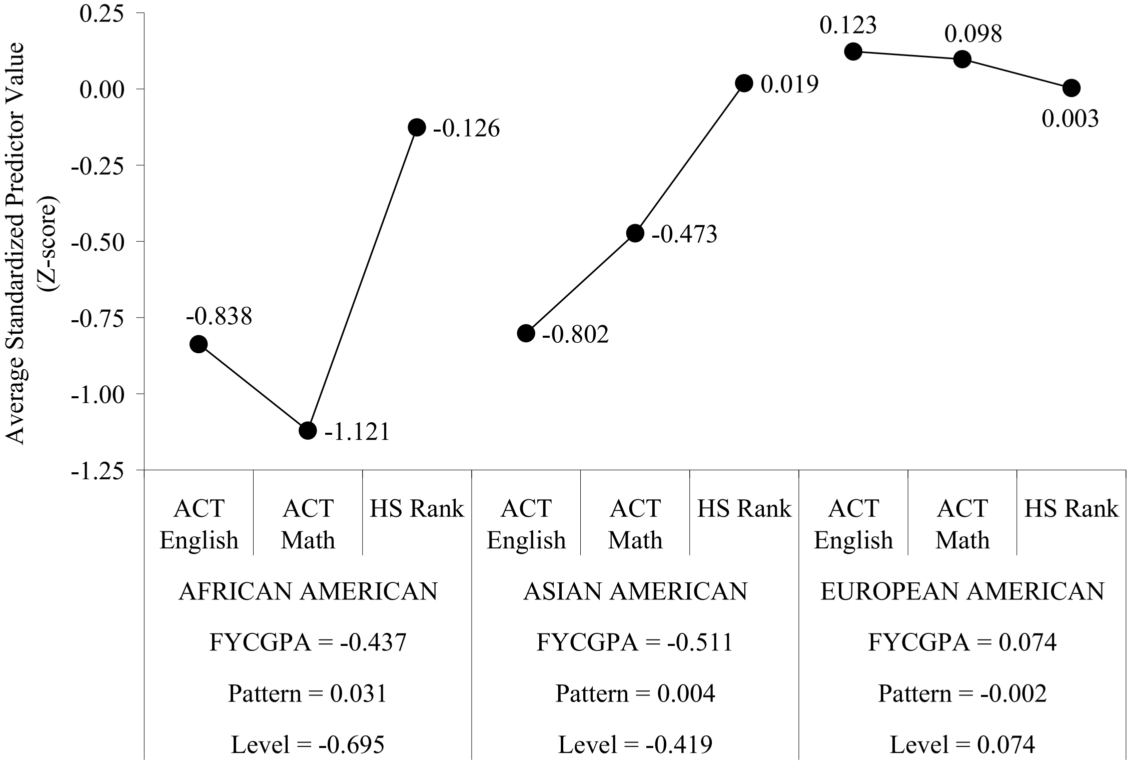

Figure 3. Average predictor profiles by race.

individual differences in both level and pattern. To answer this question we need to understand the amount of variation that was captured by level and pattern independently (note that the bivariate correlation between level and pattern equaled -0.11, so the effects were nearly orthogonal). Table 1 presents the standardized slope coefficients (later denoted as β), p-values, and part-correlations for level (β = 0.496; p < 0.001) and pattern (β = 0.138; p < 0.001). A significant F-test ($F(2, 1{,}929) = 23.6$; p < 0.001) provided evidence that pattern accounted for variation in college grades after controlling for level.

The F-test results suggest that statistically optimal weights provided some predictive value that was not captured by equal weights. Still, the results suggest that the vast majority of variance accounted for in CGPA was attributed to level, or an equal weighting scheme. Squaring the part correlations in Table 1 yields the change in R^2 effect-size. Level alone accounted for 23.1% of the total variance in CGPA or 92.5% of the explained variance in CGPA (e.g., 0.231/0.250 = 0.925). Pattern accounted for 1.9% additional variance in CGPA or 7.5% of the explained variance in CGPA. This evidence suggests that individual differences in CGPA were more associated with differences in level than with pattern. From a prediction perspective, equal weights captured most of the variance in CGPA when compared to statistically optimal weights.

Davison and Davenport (2002) note that it is also important to cross-validate CPM findings to assess the value of level and pattern in a different sample. The total sample was randomly divided into two groups. Standardized test scores and high school rank were regressed onto CGPA and the resulting standardized slope coefficients were used to create pattern variables in the omitted sample. The bottom portion of Table 1 summarizes the cross-validation findings. In particular, level accounted for 23.3% and 22.9% of the total variance in CGPA within each sample. Pattern accounted for an additional 2.6% and 1.3% after controlling for level. The cross-validation results provide additional evidence that level accounted for the vast majority of variance in CGPA. Note that this cross-validation is especially important given sample fluctuations of regression weights and the fact that practitioners may use them as if they are stable.

This finding suggests that the differences among CGPA scores for the three subjects in Figure 2 were largely due to individual differences in level rather than pattern; i.e., subjects with higher academic performance in college tended to have higher scores on all of the predictors rather than a configuration of predictor values that matched the criterion profile. For instance, Figure 2 shows that the subject who had the highest values on all of the predictors (Subject 122) also had the largest CGPA. Conversely, the two subjects that exhibited variability in their standardized predictor values had lower level values and lower academic performance.

Differential Prediction Application of the Criterion Profile Methodology

The previous section presented a profile application of the CPM. This section focuses upon the use of the CPM as a means of comparing the predictive value of statistically optimal weights and/or equal weights in differential prediction studies. Previous research suggests that an equally weighted linear composite of the independent variables provides predictive power comparable to or better than statistically optimal weights (Dana & Dawes, 2004; Davis & Sauser, 1991; Dawes, 1979; Dawes & Corrigan, 1974; Dawes et al., 1989). The use of the CPM for differential prediction studies offers a way to examine whether equal weights and/or statistically optimal weights are differentially valid for different groups simultaneously.

The CPM statistical results in Table 1 provided evidence that equally weighting the independent variables accounted for nearly all of the variation in CGPA. Thus, the value in using statistically optimal weights after controlling for an equally weighted linear composite was limited. In this instance, where level accounts for the majority of variation in CGPA, it may not be appropriate to assess differential prediction of subgroups by comparing the statistically optimal regression equations. Instead, a more meaningful differential prediction study should independently compare the extent to which level, a composite that equally weights the independent predictors, and pattern differentially predicts CGPA for students of different races/ethnicities.

Given the relative value of equal and statistically optimal weights in this study, it may be statistically appropriate to exclude pattern from the MMR model and only estimate whether level differential predicts college grades. Instead, pattern was included in the differential prediction model to demonstrate how researchers can use the CPM to address situations where level and pattern each account for a significant amount of variation in the criterion.

Table 2 presents CPM results for comparing the equivalence of subgroup regression equations. The estimated CPM-MMR model is shown below in Equation 5:

$$FYCGPA_p = b_0 + b_1\left(\mathrm{AFA}_p\right) + b_2\left(\mathrm{ASA}_p\right) + b_3\bar{X}_p + b_4 Cov_p + b_5\left(\mathrm{AFA}_p\right)\bar{X}_p + b_6\left(\mathrm{ASA}_p\right)\bar{X}_p + b_7\left(\mathrm{AFA}_p\right)Cov_p + b_8\left(\mathrm{ASA}_p\right)Cov_p + e_p \qquad (5)$$

where the dummy variables: AFA_p and ASA_p, equal one for African American and Asian American students, respectively, and zero otherwise. Additionally, CGPA, level, and pattern were standardized. b_0 represents the average EA CGPA in standard deviation units, since EA was the reference group. b_1 and b_2 represent the average difference in CGPA between AFA and EA and ASA and EA, respectively. In other words, these two parameters represent necessary intercept adjustments to better predict these two minority groups relative to the majority group. Furthermore, the EA coefficient for level was b_3 and b_5 and b_6 denote the amount that the slope coefficient for level differed between AFA and EA and ASA and EA, respectively. Just as b_1 and b_2 represented adjustments, so does b_5 and b_6. b_4 represents the contribution of pattern to the prediction equation for EA. b_7 and b_8 represent the corresponding adjustments to pattern for AFA and ASA, respectively.

Table 2. Criterion Profile Methodology Moderated Multiple Regression Summary

	Slope	Sig.
Intercept	0.031	
African American	-0.040	
Asian American	-0.181	*
Level	0.459	***
African American * Level	0.066	
Asian American * Level	0.182	*
Pattern	0.140	***
African American * Pattern	-0.015	
Asian American * Pattern	0.043	

Note. The model $R^2 = 0.259$.
Level and Pattern were standardized.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The extent to which all adjustments are non-significant indicates the degree to which one regression equation would be equally predictive for the three racial/ethnic groups. Statistically significant b_5 and/or b_6 parameters would indicate group differences in the accuracy of equal weights for predicting college grades. Similarly, statistically significant b_7 and/or b_8 parameters would suggest that the ethnic groups were not equal relative to using optimal weights as represented by the beta coefficients.

The evidence in Table 2 suggests that AFA and EA did not statistically differ in regression equations. None of the adjustments for the AFA group were significant which indicates that the estimates provided by the referent group, EA, were sufficient for AFA. That is, level ($b_5 = 0.066$; $p > 0.05$) and pattern ($b_7 = -0.015$; $p > 0.05$) predicted the same for AFA and EA and there were no intercept differences ($b_1 = -0.040$; $p > 0.05$), which indicated that the AFA and EA groups performed similarly academically after controlling for the three measures of pre-collegiate academic success. Conversely, ASA exhibited a statistically lower intercept than EA ($b_2 = -0.18$; $p < 0.05$) and a larger slope coefficient for level ($b_6 = 0.18$; $p < 0.05$). The former finding suggests that the ASA group tended to earn lower CGPA than the EA group. The slope coefficients in equation 5 represent partial correlations, so the latter finding suggests that equal weights demonstrated more criterion-related validity for ASA students than for their EA counterparts. Moreover, in additional analyses, level accounted for 22.6% and 21.0% of the variance in CGPA for ASA and EA, respectively, which suggests that equal weights may have been slightly more accurate for ASA students. EA and ASA did not differ in the extent to which using statistically optimal weights related to the subsequent quality of the prediction.

Conclusion

This paper used the CPM to: 1) conduct profile analysis to differentiate between individuals and groups who earn high and low college grades; and 2) explore differential predictability in the use of equal and statistically optimal weights. The results provided evidence that equal weights, or level, accounted for more variance in CGPA than statistically optimal weights, pattern. The criterion profile, or statistically optimal weights, provided little additional predictive ability for differentiating between students with high and low CGPA scores. Therefore, the best weighting scheme was one that treated test scores and high school rank equally rather than the statistically optimal weighting scheme that gave more weight to ACT English and high school rank than to ACT Math. Perhaps intuition or research would suggest that ACT Math is not as good of a predictor of academic performance for students in liberal arts, since their coursework may include less mathematics. The findings of this study suggest that using equal weights for all applicants will capture most of the variability in first-year grades; approximately 90% of the accounted for variance. From a practitioner’s perspective, the evidence suggests that ACT Math scores should be treated with the same weight or importance as ACT English scores, and high school rank in decisions for admitting applicants to the College of Liberal Arts.

This finding had direct relevance for assessing differential prediction. That is, statistically optimal weights accounted for very little variance in addition to equal weights. Thus, it was more appropriate to test the extent to which equal and statistically optimal weights differentially predicted grades for different racial groups. In fact, an equal weighting scheme was more valid for ASA than for EA. There was no evidence to suggest that AFA and EA equations differed, so the equal and statistically optimal weighting schemes provided similar predictive accuracy.

The use of equal or statistically optimal weights poses another methodological challenge for assessing differential prediction. Future differential prediction studies should determine the value of equal and statistically optimal weights by computing the variance accounted for by level and pattern in college grades in the full sample and in cross-validation samples. Failure to determine the relative value of equal and statistically optimal weights may result in researchers comparing the equivalence of groups in statistically optimal coefficients when equal weights account for the vast majority of variance in a criterion. This study demonstrated that the CPM is an appropriate method for addressing this methodological issue.

References

Breland, H. M. (1979). *Population validity and college entrance measures* (Research Monograph No. 8). New York: College Entrance Examination Board.

Bridgeman, B., McCamley-Jenkins, L., & Ervin, N. (2000). *Predictions of freshman grade-point average from the revised and recentered SAT I: Reasoning Test* (Research Report No. 2000-1). New York: College Entrance Examination Board.

Burton, N. W., & Ramist, L. (2001). *Predicting success in college: SAT® studies of classes graduating since 1980* (Research Report No. 2001-1). New York: College Entrance Examination Board.

Chou, T., & Huberty, C. J. (1990). *A freshman admissions prediction equation: An evaluation and recommendation* (No. ERIC Document Reproduction Service No. ED 333 081). Athens, Georgia: University of Georgia.

Cleary, T. A. (1968). Test bias: Prediction of grades of Negro and white students in integrated colleges. *Journal of Educational Measurement, 5*, 115-124.

Cleary, T. A., Humphreys, L. G., Kendrick, S. A., & Wesman, A. (1975). Educational uses of tests with disadvantaged students. *American Psychologist, 30*, 15-41.

Dana, J., & Dawes, R. M. (2004). The superiority of simple alternatives to regression for social science predictions. *Journal of Educational and Behavioral Statistics, 29*, 317-331.

Davis, K. R., & Sauser, W. I. (1991). Effects of alternative weighting methods in a policy-capturing approach to job evaluation: A review and empirical investigation. *Personnel Psychology, 44*, 85-127.

Davison, M. L., & Davenport, E. C., Jr. (2002). Identifying criterion-related patterns of predictor scores using multiple regression. *Psychological Methods, 7*, 468-484.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34*, 571-582.

Dawes, R., M. & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81*, 95-106.

Dawes, R., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243*, 1668-1674.

Dorans, N. J. (2004). Using subpopulation invariance to assess test score equity. *Journal of Educational Measurement, 41*, 43-68.

Duran, R. P. (1983). *Hispanics' education and background: Predictors of college achievement* (Research Report). New York: College Entrance Examination Board.

Einhorn, H. J., & Hogarth, R. M. (1975). Unit weighting schemes for decision making. *Organizational Behavior and Human Performance, 13*, 171-192.

Elliott, R., & Strenta, A. C. (1988). Effects of improving the reliability of the GPA on prediction generally and on comparative predictions for gender and race particularly. *Journal of Educational Measurement, 25*, 333-347.

Hills, J. R. (1964). Prediction of college grades for all public colleges of a state. *Journal of Educational Measurement, 1*, 155-159.

Humphreys, L. (1952). Individual differences. *Annual Review of Psychology, 3*, 131-150.

Linn, R. L. (1973). Fair test use in selection. *Review of Educational Research, 43*, 139-161.

McKelpin, J. P. (1965). Some implications of the intellectual characteristics of freshmen entering a liberal arts college. *Journal of Educational Measurement, 2*, 161-166.

Munday, L. (1965). Predicting college grades in predominately negro colleges. *Journal of Educational Measurement, 2*, 157-159.

Noble, J. (2003). *The effects of using ACT composite score and high school average on college admission decisions for racial/ethnic groups* (ACT Research Report Series No. 1). Iowa City: ACT Program.

Pennock-Román, M. (1994). *College major and gender differences in the prediction of college grades* (College Board Report No. 2). New York: College Entrance Examination Board.

Ramist, L., Lewis, C., & McCamley-Jenkins, L. (1993). *Student group differences in predicting college grades: Sex, language, and ethnic groups* (College Board Report No. 1). New York: College Entrance Examination Board.

Richards, J. M., Jr., & Lutz, S. W. (1968). Predicting student accomplishment in college from the ACT assessment. *Journal of Educational Measurement, 5*, 17-29.

Sassenrath, J. M., & Pugh, R. (1965). Relationships among CEEB Scholastic Aptitude Test and American College test scores and grade point average. *Journal of Educational Measurement, 2*, 199-205.

Saunders, D. R. (1956). Moderator variables in prediction. *Educational and Psychological Measurement, 16*, 209-222.

Stanley, J. C., & Porter, A. C. (1967). Correlation of scholastic aptitude test scores with college grades for negroes versus whites. *Journal of Educational Measurement, 4*, 199-218.

Watkins, M. W. (2000). Cognitive profile analysis: A shared professional myth. *School Psychology Quarterly, 15*, 465-479.

Wilson, K. M. (1983). *A review of research on the prediction of academic performance after the freshman year* (College Board Research Report No. 83-2 and Educational Testing Service Research Report No. 83-11). New York: College Entrance Examination Board.

Young, J. W. (1994). Differential prediction of college grades by gender and by ethnicity: A replication study. *Educational and Psychological Measurement, 54*, 1022-1029.

Young, J. W. (2001). *Differential validity, differential prediction, and college admission testing: A comprehensive review and analysis* (Research Report No. 2001-6). New York: College Entrance Examination Board.

Send correspondence to: Steven Andrew Culpepper
University of Colorado Denver
Email: steve.culpepper@ucdenver.edu