DOCUMENT RESUME

ED 104 182 HE 006 284

AUTHOR Cimbala, S. J.
TITLE Campus Location and Student Performance: A Preliminary Examination.
PUB DATE 3 Dec 74
NOTE 16p.

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS *Academic Achievement; Campuses; *Higher Education; *Low Achievement Factors; Performance Factors; Predictive Measurement; *School Environment; Statistical Data; *Student Experience
IDENTIFIERS *Pennsylvania State University

ABSTRACT

Several studies have attempted to hypothesize about the probable effects on student academic performance at the Commonwealth campus versus University Park freshman admission for the Pennsylvania State University. This report imposes a framework of predictive modelling on the data in order to derive some notion of the weight of variables in predicting academic success in the university. Within that context, the importance of the issue of campus location at entry can be examined. The author concludes that whether students begin their careers at Penn State at the Delaware County Campus or at University Park, based on inferences that can be made from the data presented, makes little or no difference when effects are measured in multivariate procedures. The statistical tables included in the text are reproduced from the best available copy. (Author/PG)

CAMPUS LOCATION AND STUDENT PERFORMANCE:

A Preliminary Examination*

S. J. Cimbala
Office of Academic Affairs
Delaware County Campus
Pennsylvania State University

December 3, 1974

* Acknowledgments are due to Edmond Marks, Office of Budget and Planning, for providing data used in this study, and to Thomas Smith, a student at the Delaware County Campus, for assistance in data processing.

I. APPROACH AND ASSUMPTIONS

Several studies have attempted to hypothesize about the probable effects on student academic performance of Commonwealth campus versus University Park freshman admission. This question can be conceptualized in various ways, and the approach that follows does not preclude alternatives. Nevertheless, it seems that some issues are theoretically and logically prior to others, and a descriptive mapping of the data through multivariate analysis can disentangle some of these at the outset. What follows is an effort to impose the framework of predictive modelling upon the data, in order to derive some notion of the relative "weight" of various variables in predicting academic success in the university. Within that context, the importance of the issue of campus location at entry can be examined. Data for the freshman class of 1968 at University Park and at the Delaware County Campus will be used in the study.[1]

The question whether "campus of first admission" makes a difference for students' eventual performances is fundamentally a question about the relative strength of relationship among theoretically significant variables and about the relative predictive power of those variables. Campus location

[1] Data were provided by Edmond Marks of the Office of Budget and Planning.

variables are either more or less informative than other variables; their importance cannot be presumed or discussed in isolation. Thus, the theoretical problem suggests multivariate procedures as appropriate explanatory and predictive strategies for analysis. In the case that follows, models are generated through multiple regression analysis in which the hypothesized dependent variable, ultimate grade point average, is predicted from various combinations of independent variables. These combinations of independent variables are grouped within cognate models of academic, demographic and "mixed" predictor variables. Academic, demographic and mixed models are first compared for all students in the data set, then for the subset of students who graduated, and finally for the subset of students who withdrew prior to graduation. Thus, nine regression models will be generated for the entire analysis: three "demographic models," three "academic" models and three "mixed" models.

The assessment of the relative importance of campus location at admission, within the multivariate models, can be facilitated by "Stepwise" multiple regression procedures, which build predictor variables into the equation in rank order of relative weight*. This strategy allows

* Reflected in Table III, Summary of R^2 Change Values

evaluation of the predictive importance of campus location after controlling for other possibly significant variables, such as high school average, SAT scores, demographic variables, and so forth. This way of looking at the problem avoids confusion of the campus effect per se with the effects of other variables which are correlated with both campus effect and student performances. Since our University admissions policies during this period tracked students to campuses on the basis of high school GPA (as well as other factors), it is logical to suspect the existence of these multicollinearities which might confound the analysis.

Another advantage of these procedures is the comparative perspective on the predictive and explanatory power of academic and demographic variables, and on their possible interrelations as predictors in the same model. We know from earlier studies in the literature that measures of high school performance, verbal and mathematical aptitude scores, socioeconomic backgrounds of students and other variables are likely to present problems of multicollinearity and interaction effects when used as simultaneous predictors. In the "real world" these measures are confounded by the same problem, of overlapping measurement of shared attributes among supposedly discrete phenomena. To get at these possible difficulties, we propose the following decomposition of prediction equations: (1), in the "academic" regression models, demographic variables are omitted from the equation and "campus of origin" is compared with academic predictors only; (2), in the "demographic" regression models, academic variables are omitted from the analysis and "campus of origin" is compared with demographic predictors only; (3), the final step involves the comparisons of campus of origin with both academic and demographic predictors within the same model. The academic-location comparisons, demographic-location comparisons and the "mixed" comparisons will be run for each of three groups: all students in the data set, those who graduated only, and those who withdrew from the university before graduation.

In interpreting the results of these regression models, the rules of thumb for various statistical indicators will be as follows. Zero order product moment correlation coefficients (Pearson R) estimate the strength of relationship (linear) between two variables. The squared product moment correlation coefficient (R^2) estimates the proportion of variation in one variable "explained" by or accounted for by the second variable. Partial correlation coefficients estimate the strength of relationship between variables "controlling" for the effects of other variables in the model. The slope of the linear least squares regression line is given by the regression coefficient (b) and the relative predictor weight for variables in a multiple regression equation by the standardized regression coefficient B (Beta). Variables have been recoded in some cases to treat all measurements as interval for analytic purposes.

Variables in the analysis were as follows. For the academic model, High School Average (HSA), High School Rank (HSR), Scholastic Aptitude Test-Verbal, Mathematical, and Total (SATV, SATM, SATT) were used as predictors in addition to Campus of Admission (ADMCAMP). In the demographic model, Father's Education (FATHEDUC), Father's Occupation (FATHOCC), Income (INCOME) and Campus of Admission were used as predictors. All variables were used as predictors in the mixed model. The number of cases for the models were 2774 for all students, 1899 for those who graduated and 852 for those who withdrew prior to graduation. (Discrepancies may be accounted for by difficulties in unambiguous classification or missing data. Coding criteria were determined by OBP. Raw data output with descriptive statistics not reported in this paper may be obtained from the investigator).

TABLE I. Regression Coefficients for Predictors of Final Cumulative GPA, by Model Type and Student Subgroup

	ALL STUDENTS			GRADUATED			WITHDREW		
	ACA	DEM	ACA+DEM	ACA	DEM	ACA+DEM	ACA	DEM	ACA+DEM
ADMCAMP	-.06132	-.45192	-.03277	.07486	-.21305	.08992	.20258	-.23913	.23440
FATHEDUC		.01568	.01251		.01186	.00916		.01858	.01510
FATHOCC		-.00060	-.00847		-.00120	-.00242		-.00585	-.00264
HSA	.49265		.49334	.36323		.36419	.54727		.54635
HSR	-.04339		-.05304	-.01996		-.02951	-.03058		-.03837
INCOME		.01418	.01452		.00919	.01110		.01629	.00665
SATM	-.00051		-.00047				.00044		.00041
SATT	.00038		.00087	.00040		.00034			
SATV				.00059		.00057	.00165		.00152
CONSTANT	.54459	2.98719	.43149	.99090	3.00523	.91335	-.65071	2.30751	-.76674

SUMMARY OF B VALUES WITH CONSTANTS.

TABLE II. "STANDARDIZED" REGRESSION COEFFICIENTS (BETA WEIGHTS) FOR PREDICTORS OF FINAL CUMULATIVE GRADE POINT AVERAGE, BY MODEL TYPE AND STUDENT SUBGROUP

	ALL STUDENTS			GRADUATED			WITHDREW		
	ACA	DEM	ACA+DEM	ACA	DEM	ACA+DEM	ACA	DEM	ACA+DEM
ADMCAMP	-.02300	-.16969	-.01229	.02956	-.08410	.03551	.09025	-.10661	.10443
FATHEDUC		.08839	.07041		.09333	.07178		.09367	.07636
FATHOCC		-.00830	-.01610		-.00777	-.01616		-.02346	-.02402
HSA	.38151		.38204	.38446		.38547	.38080		.38315
HSR	-.04530		-.05537	-.02601		-.03846	-.03894		-.04134
INCOME		.03078	.03149		.02932	.03412		.03673	.01265
SATM	-.02357		-.02798				.04877		.04517
SATT	.20843		.18511	.11907		.10197			
SATV				.11030		[illegible]	[illegible]		.17645

Summary Table of BETA VALUES.

TABLE III. SUMMARY OF R^2 CHANGE VALUES FOR PREDICTORS OF FINAL CUMULATIVE GPA, BY MODEL TYPE AND STUDENT SUBGROUP

	ALL STUDENTS			GRADUATED			WITHDREW		
	ACA	DEM	ACA+DEM	ACA	DEM	ACA+DEM	ACA	DEM	ACA+DEM
ADMCAMP	.00043	.03495	.00012	.00068	.00729	.00097	.00389	.01090	.00567
FATHEDUC		.01645	.00793		.01375	.00886		.01979	.00594
FATHOCC		.00001	.00017		.00004	.00017		.00043	.00059
HSA	.21015		.21015	.20431		.20431	.18786		.18786
HSR	.00145		.00185	.00042		.00087	.00069		.00096
INCOME		.00075	.00102		.00075	.00097		.00097	.00012
SATM	.00121		.00110				.0017[illegible]		.00158
SATT	.02505		.02505	.00321		.00225			
SATV				.03879		.03879	.03111		.03111

Summary of R square Change Values.

TABLE IV. VALUES OF SIMPLE R FOR PREDICTORS OF FINAL CUMULATIVE GPA, BY MODEL TYPE AND STUDENT SUBGROUP.

	ALL STUDENTS			GRADUATED			WITHDREW		
	ACA	DEM.	ACA+DEM	ACA	DEM	ACA+DEM	ACA	DEM	ACA+DEM.
ADMCAMP	-.19333	-.18696	-.19333	-.09559	-.09566	-.09559	-.12660	-.13256	-.12660
FATHEDUC		.12911	.12977		.11726	.12011		.14068	.13791
FATHOCC		-.06335	-.06035		-.07191	-.07111		-.07279	-.06658
HSA	.45842		.45842	.45201		.45201	.43343		.43343
HSR	-.31145		-.31145	-.28933		-.28933	-.27055		-.27055
INCOME		.09371	.09324		.08030	.08211		.10173	.09752
SAT M	.20586		.20586				.19370		.19370
SAT T	.27701		.27701	.30229		.30229			
SAT V				.29730		.29730	.27748		.27748

Summary of Simple R values.

TABLE V. SUMMARY OF MULTIPLE R VALUES FOR PREDICTORS OF FINAL CUMULATIVE GPA, BY MODEL TYPE AND STUDENT SUBGROUP

	ALL STUDENTS			GRADUATED			WITHDREW		
	ACA	DEM	ACADEM	ACA	DEM	ACADEM	ACA	DEM	ACADEM
ADMCAMP	.48814	.13696	.49738	.49699	.14506	.50612	.47208	.17519	.48019
FATHEDUC		.21309	.49308		.11726	.50197		.14068	.47425
FATHOCC		.21493	.49726		.14777	.50715		.17914	.48343
HSA	.45848		.45842	.45201		.45201	.43343		.43343
HSR	.48646		.49496	.49741		.50698	.47466		.48271
INCOME		.21492	.49709		.14763	.50516		.17794	.48355
SATM	.48771		.49607				.47395		.48183
SATT	.48497		.48497	.49630		.50420			
SATV				.49306		.49306	.46795		.46795

SUMMARY OF MULTIPLE R VALUES

TABLE VI. SUMMARY OF R^2 VALUES FOR PREDICTORS OF FINAL CUMULATIVE GPA, BY MODEL TYPE AND STUDENT SUBGROUP

	ALL STUDENTS			GRADUATED			WITHDREW		
	ACA	DEM	ACA+DEM	ACA	DEM	ACA+DEM	ACA	DEM	ACA+DEM
ADMCAMP	.23828	.03495	.24739	.24700	.02104	.25615	.22286	.03069	.23058
FATHEDUC		.04541	.24313		.01375	.25197		.01979	.22491
FATHOCC		.04620	.24727		.02184	.25720		.03269	.23370
HSA	.21015		.21015	.20431		.20431	.17786		.17786
HSR	.23665		.24498	.24741		.25703	.22531		.23311
INCOME		.04619	.24710		.02179	.25519		.03166	.23382
SATM	.23786		.24608				.22463		.23215
SATT	.23520		.23520	.24631		.25422			
SATV				.24311		.24311	.21897		.21897

Summary of R Square Values.

II. CONCLUSIONS

1. High school average is the strongest relative predictor of academic success measured as ultimate grade point average in the university, for students in general and for the "graduated" and "withdrew subsets.

2. Other variables, whether demographic or academic, add very little to the explanatory power (R^2) gained from knowledge of high school average alone. SAT scores improve the R^2 modestly in the academic model, but the additive effects of other variables in other models are modest to insignificant.

3. "Campus of first admission" has more relative clout in explaining or predicting academic success when grouped with demographic variables alone, as opposed to its utility in the academic or mixed model. This is a function of the weakness of the demographic variables, however. All demographic variables are relatively weak predictors of academic success, as is campus location, when their effects are controlled for academic variables.

4. The effects of campus location at entry are somewhat more important for students who ultimately withdrew prior to graduation than for students

who graduated, but the differences are not significant. (All the coefficients are statistically significant due to large size of the "sample," but substantive significance is interpretable through the relative size of R^2 and B(Beta)).

5. In the mixed and academic models, the partial correlation between campus location and academic success is weak to nonexistent. Taking into account academic variables which are correlated in the real world (and in our admissions policy) with campus location, we reduce the explanatory power and predictive utility of campus location variables to virtually no relationship.

Summary

Whether students begin their careers at Penn State at the Delaware County Campus or at University Park, based upon inferences that can be made from these data, makes little to no difference when effects are measured in multivariate procedures as described above. This does not mean that other procedures cannot reveal trends and patterns that show otherwise, for the outcomes of research procedures are derivatives of the assumptions (about the universe of investigation and the form of the data) inherent in the selection of research strategy. The reader is directed to the selection of research strategy as it relates to the questions of interest to the investigator, as stated.