

Measuring the Success of Transfer Students: A New Metric Evaluated

Gerald W. McLaughlin, Emeritus, DePaul University

Jacqueline E. McLaughlin, University of North Carolina – Chapel Hill

Josetta S. McLaughlin, Emerita, Roosevelt University

Richard D. Howard, Higher Education Consultant

Sandra Whalen, CSRDE

Presented at

2016 AIR Forum

New Orleans, LA

June 1, 2016

Measuring the Success of Transfer Students: A New Metric Evaluated¹

Gerald W. McLaughlin, Jacqueline E. McLaughlin, Joesetta S. McLaughlin, Richard D. Howard, and Sandra Whalen

U.S. institutions of higher education are regularly evaluated by the success of their students as measured by the retention and graduation rates of a first-time, full-time, degree-seeking cohort. A number of concerns have been raised about use of this type of cohort to create and interpret indicators of student success as a means for defining institutional performance. The data, provided to the federal government by colleges or universities, are typically used to calculate student retention and graduation and represent only the graduation of a beginning cohort of students (i.e., the success of first-time, full-time beginning students noted above) at a specific institution. Critics argue that this approach is fundamentally flawed and that it creates an incomplete and inaccurate picture of what is happening at U.S. colleges and universities. Among other criticisms, they point out that the measure used is inappropriate for four-year institutions where community colleges are increasingly used as pathways to completion. Furthermore, they note that the measure is totally inappropriate for institutions where the traditional cohort is a small proportion of the beginning enrollments (for example, in urban institutions where many students enter as part-time).

The purpose of this research is to address critics' concerns about the current use of only first-time, full-time, degree-seeking students to calculate retention and graduation rates as a measure of student success, especially for colleges and universities with populations of transfer students from community colleges. We demonstrate the use of a methodology that incorporates an additional cohort (the community college transfer) and demonstrate that the resulting statistic based on a redefined additional cohort—those completing 30 hours at a community college—is sufficiently reliable and valid². The additional cohort defines transfers as those completing 30 hours at a community college. These data were reported over a

¹ This paper reports results from an extension of research presented as Looking at the Success of Community College Transfer Cohorts with CSRDE Data: An Exploratory Analysis, National Symposium on Student Retention, Orlando, Florida. November 2015.

² A different methodology --The Student Achievement Measure -- has also been proposed as an alternative but uses a different definition of transfers.

ten-year period to the Consortium for Student Retention Data Exchange (CSRDE) by 117 senior institutions with cohorts detailed by gender and by ethnicity.

The data are investigated for reliability and validity when used in describing institutional outcomes on student success. We first examine the magnitude of the measures, their statistical stability and how they correlate between and within themselves. The correlations of a select group of the measures with institutional characteristics are then examined. Next, we examine the ability of the data to support meaningful interpretation from a two level hierarchical linear model. The intent is to show that if these methods produce interpretable results, then the metric is sufficiently reliable and valid to be pursued to complement the IPEDS defined cohort. Finally, the statistical issue around interpreting the institutional characteristics as explaining student success is discussed. The predictors of average institutional outcomes are not necessarily the same as the predictors of individual student outcomes. While various measures may help administrators understand the institutional level measure of student success, the measures cannot be assumed to explain the success of individual students. This problem is known as the ecological fallacy and has been well documented in behavioral science methodology (Trochim, 2006). This is discussed in more detail in Appendix A-2.

Background

Student success, typically measured by retention and graduation of first-time, full-time, degree-seeking students, is one of the key performance indicators used in management of colleges and universities. However, its use does not stop here. The construct is also used in various ratings and rankings of U.S. colleges and universities such as those developed by *U.S. News and World Report*. It is a core component in political discussions about accountability and is found in various systems designed to improve the higher education experience for students and the understanding of alternatives by students and their parents. Funding agencies that provide millions of dollars with the stated intent to improve the graduation of students use it as a key construct. State legislatures look at the educational pathway, trying to determine where there are effective and efficient transitions, often from high school to community college to senior institution. A good example of this use is the initiative Tennessee Transfer Pathways which is designed to facilitate the transfer from community college to a senior institution (See <http://www.tntransferpathway.org/>, n.d.). There are also organizations with websites funded by various grants that make student success at both two-year and four-year institutions available to the public. (See

<http://www.collegemeasures.org> -- funded by the Gates Foundation). Finally, the term “student success” is used in statements of concern by various advocacy groups that focus on those who are seen as disadvantaged in our society. Unfortunately, the only graduation and retention statistics nationally available are computed on cohorts that specifically exclude community college transfers.

Almost all of these groups and voices indicate their understanding that institutional characteristics—and particularly the characteristics of students who enter the institution—heavily influence the results of graduation and retention rates.³ Statistical evidence that institutional characteristics are associated with student graduation is shown by *U.S. News and World Report*’s regression model which predicts graduation based on institutional characteristics. *U.S. News and World Report* then extends this concept to identify under- and over-achieving institutions. Scholarly research does confirm that a regression model frequently fits the average outcome of various measures of student success (Pike & Robbins, 2015; Velazquez, 2015).⁴ There are, however, statistical and data problems associated with interpreting the institutional characteristics as explaining student success. The statistical issue is that the predictors of average institutional outcomes are not necessarily the same as the predictors of individual student outcomes. As noted earlier, this problem is known as the ecological fallacy and has been well documented in behavioral science methodology (Trochim, 2006). Again, the ecological fallacy is an important interpretation issue when relating the predictors of average institutional outcomes to the predictors of individual student outcomes—they often are not the same. Valid conclusions cannot be drawn about the success of individuals from analyses using average institutional level data. Both types of predictor data—institutional and individual-level data—are important for different purposes. Political and policy questions typically focus at the institutional level and are typically answered by average institutional level data. In general, the desire is to increase the overall average graduation or retention ratios. Student level predictors on the other hand are typically used by practitioners who are looking for correlates that are related to graduation or retention. The ecological fallacy is discussed in further detail in Appendix A-2.

³ Several examples of these statements are included in the Appendix to this paper to include extracts from an article from *Inside Higher Education* and an extract from the Department of Education and NPEC.

⁴ Also see “Graduation Rates and Student Success: Squaring Means and Ends”, *Perspectives*, American Association of State Colleges and Universities, Fall 2006, retrieved from http://www.aascu.org/uploadedFiles/AASCU/Content/Root/PolicyAndAdvocacy/PolicyPublications/06b_perspectives%2081%29.pdf.

Measuring the Success of Transfer Students

The data problem occurs because the success of students being modeled in most analyses is almost always the students that are included in the IPEDS cohort. These are students who enroll in an institution as first-time, full-time, degree-seeking students. For some institutions, this is a good representation of their student body; for other institutions, this is an extremely poor representation of their student body. In the latter case, the transfer cohort may be as large as, or larger than the first-time student cohort.⁵ Furthermore in no way does use of the IPEDS cohort accurately represent the effectiveness of pathway programs.

A recent initiative to address the problem of transfer cohort size is the Student Achievement Model (SAM, n.d.). As explained in Appendix A-6, this is a bachelor's seeking model that looks at full-time and part-time entering and transfer students. While this is more inclusive than the IPEDS definition, it possibly is overly inclusive for some uses such as the value of community college transfer programs. It allows the analysis of transfer students, but there is no consideration of whether they came from a community college or a four-year institution. The result is that it cannot address issues of the transfer gap between community colleges and senior institutions. Also, there is no differentiation of the intent of the students to take courses. They may have had one course at another institution and are just shopping around or they may have had several years of study and represent very serious students.

A second initiative designed to address the inclusion of transfer students in student success statistics came from CSRDE. In 2005, members of CSRDE voiced a need to evaluate the success of transfer students within their institutions. The challenge in developing these data was the tremendous variation in characteristics of transfer students. Some came from two-year institutions, some came from four-year institutions, and some had simply taken several courses at a two-year institution while waiting to enter a four-year institution. Some students intended to earn a degree and some were only interested in an occasional course or two. Discussions with members of the CSRDE data exchange revealed that the main concern at the institutions for transfer students involved those who had transferred from two-year institutions. A data-sharing protocol was developed with the cohort defined as the group of students entering in the fall who had completed the equivalent of at least one year at a community college.⁶ This was interpreted as a measure of those who were attending the community college as a substantial step in

⁵ A full definition of the IPEDS variables is available on the web. <http://nces.ed.gov/ipeds/glossary/>

⁶ Appendix 1 provides details about the development of the survey, why transfers from community colleges were selected for this project, other variables included in the report, and information about the CSRDE.

their intent to pursue higher education and eliminated those who incidentally took a course or two from a community college.⁷

Methods and Analysis

If the definition of a cohort is to have value it must have reliability and validity. This research examines psychometric characteristics when the definition of community college transfers with at least one year of credits is examined for ethnic cohorts of Asians, Blacks, Hispanics, and Whites. The research is important for a number of reasons. First, the usability of the data is particularly important since the data reflect a critical topic in the student success discussion—the movement from the community college to the senior institution. Second, the analyses using CSRDE data help determine whether the data on transfer students has reliability and validity similar to those data collected by IPEDS on first-time, full-time, degree-seeking students. If they do, this also demonstrates the usability of the CSRDE definitions and demonstrates the development of a data set that can be used to explore and make initial discoveries on this previously unidentified and unique cohort.

The data sample used in the following analyses is a voluntary subset of data provided by 117 senior institutions. The specific data reflect cohorts of students who have completed the equivalent of at least one academic year from a community college and who enrolled as a transfer student in the senior institution. These cohorts are tracked for seven years with the proportion graduating in each year and the proportion enrolling in the following fall term included in the data set. In addition, the percentage of the cohort enrolling initially as part-time students and the total number in the cohort are included in the data submitted to CSRDE by ethnicity.

The likely value of these data was assessed using three basic analytic methodologies that are frequently used as the bases for discussions of student success—averages, correlation, and Hierarchical Linear Modeling. The first methodology—the calculation of means—analyzes the average proportion of students graduating and the average proportion continuing enrollment. The second methodology—the use of correlations—correlates the measures of success between themselves and also with a set of student and institutional characteristics. These institutional characteristics are metrics that were

⁷ It was realized that this did not include all transfers from community colleges, but it was a balance of being inclusive and building a standardized database.

previously developed as a balanced scorecard to form peer groups and to benchmark institutional outcomes (McLaughlin, McLaughlin, & Howard, 2013). The third analysis—Hierarchical Linear Modeling (HLM) —explains variance in graduation and retention rates using the same balanced set of measures and percentage part-time by ethnicity. This replicates and advances the modeling typically used to explain the relative differences in student outcomes based on student characteristics and institutional aspects. For this situation, it is used at the institutional level with ethnicities nested within institutions.

Characteristics of the Data

The retention and graduation data used in the analyses, provided by 117 senior institutions, include cohorts of community college transfer students enrolled in the senior institution and tracked for seven years. While CSRDE participating member institutions report continuing and graduation rates by gender and also by race/ethnicity, this analysis looks primarily at the data for the combined groups of transfer cohorts from 2006 to 2012. This analysis averaged the rates over the years of the cohorts to get a more stable estimate of “after one year”, “after two years”, etc. In other words, “after one year” is for the 2006 cohort as reported for 2007, the 2007 cohort as reported for 2008, etc. The graduation within five years was selected for some of the correlation analyses and for the HLM analyses. A five-year graduation rate for the student who transfers into a four-year institution after at least one year of study at a community college would be essentially equivalent to a six-year graduation rate for a student who entered their institution as a first time/beginning freshman. This was done to produce a more stable statistic for the progression of the students. As noted earlier, the purpose of these analyses is to determine if the data collected by CSRDE on transfer students has reliability and validity similar to those data collected by IPEDS on first-time, full-time, degree-seeking students. If they do, it demonstrates the usability of the CSRDE definitions and also the development of the data set that can be used to make initial discoveries on this previously missing and unique cohort.

Characteristics of the Sample

Of the approximately 400 institutional members of CSRDE, 117 provided data on the persistence and graduation of transfers at the senior institution. The basic characteristics of this group are shown in Table 1.

Table 1: Demographic Samples of CSRDE Institutions Participating in the CSRDE Transfer Student Success Data Exchange

Urbanicity			Region		
	City	52.99%		New England/Mid-East	19.66%
	Suburb	25.64%		Plains/Great Lakes	27.35%
	Town	19.66%		Southeast/South West	17.09%
	Rural	1.71%		Rocky Mountains/Far West	34.19%
Size			Degree Offered		
	1,001-5,000	11.87%		Doctorate	57.27%
	5,001-10,000	30.51%		Masters	99.15%
	10,001-30,000	38.98%		Bachelors	100.00%
	> 30,000	18.64%		Associate	21.37%

It can be seen that these institutions are more likely to be larger institutions with enrollments between 5,000 and 30,000 students. A majority are located in cities, are from the Rocky Mountains and the West, and 99% offer a Master’s degree.

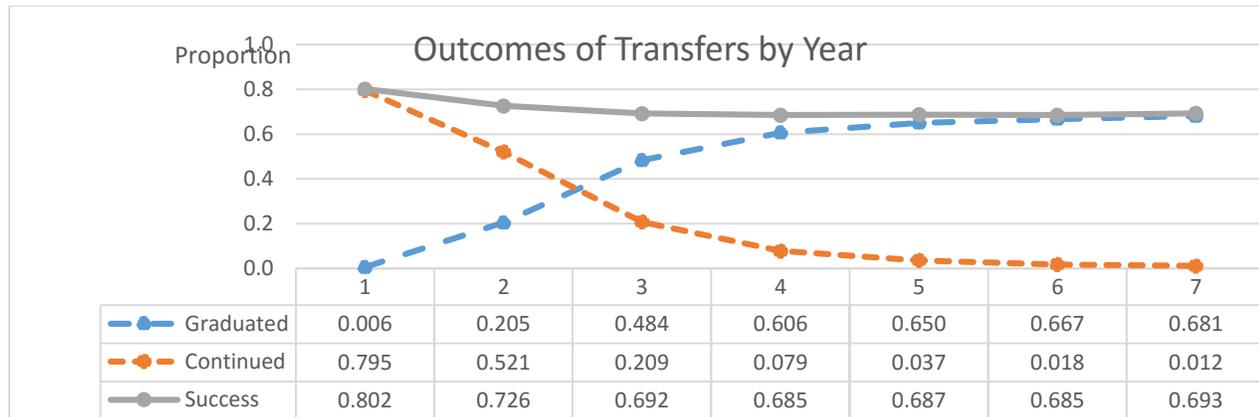
Results

Methodology 1: Means

The means for the transfers by year are shown in Figure 1. As illustrated in this figure, after one year, proportionally .006 of the transfers had graduated, and .795 continued into their second year. This gave a combined success proportion of .802 or slightly over 80%. After two years, one in five had graduated and about half (.502) continued into the third year for a combined success rate of .726 or about three-

Measuring the Success of Transfer Students

fourths. The success rate stays essentially constant at about .7 while the proportion of graduates stabilizes and the proportion of continuing students becomes very small after four to five years.⁸



Graduated: “1” indicates Graduation within 1 year; “2” indicates Graduation within 2 years; etc.

Continued: “1” indicates Continued into the second year; “2” indicates Continued into the third year; etc.

Success: “1” indicates Graduated within 1 year or Continued into the following year; etc.

Figure 1: Outcomes of Transfers by Year

The outcomes for the various ethnicities are shown in Table 2. These and the following results are for ethnic cohorts that were greater than an average of two students per year. This was done since preliminary results indicated that the very small cohorts yielded unstable results.

⁸ A recent report found that about 60% of transfers from community colleges graduated and about 10% were still enrolled four years after transfer. National Student Clearinghouse Research Center, Snapshot Report, Outcomes of Students who Transferred from Two-Year to Four-Year Institutions (Four Years After Transfer), 2012. <http://research.studentclearinghouse.org>

Table 2: Outcomes of Transfers by Year for the Four Major Ethnic Groups and Percentage of Each that Entered as Part-time

		One Year	Two Years	Three Years	Four Years	Five Years	Six Years	Seven Years	% Part-time
Asian > 2 per cohort									20.27
	Graduate	0.005	0.212	0.519	0.640	0.688	0.705	0.714	
	Continue	0.826	0.546	0.212	0.082	0.036	0.023	0.014	
	Success	0.831	0.757	0.730	0.722	0.724	0.728	0.728	
Black > 2 per cohort									21.47
	Graduate	0.006	0.158	0.390	0.511	0.556	0.583	0.598	
	Continue	0.761	0.520	0.242	0.105	0.056	0.025	0.018	
	Success	0.767	0.677	0.632	0.616	0.612	0.609	0.615	
Hispanic > 2 per cohort									21.43
	Graduate	0.004	0.207	0.488	0.608	0.653	0.672	0.692	
	Continue	0.810	0.533	0.219	0.085	0.044	0.026	0.011	
	Success	0.815	0.741	0.707	0.693	0.697	0.698	0.703	
White > 2 per cohort									18.94
	Graduate	0.006	0.220	0.508	0.632	0.676	0.691	0.702	
	Continue	0.804	0.518	0.200	0.073	0.031	0.017	0.010	
	Success	0.810	0.738	0.708	0.705	0.708	0.709	0.712	

Methodology 2: Correlation Analysis

Overview

In the following section the correlation of the success measures of transfers and beginning freshmen with other institutional characteristics is examined. The purpose of the analysis presented in this section is to provide some insight into the reliability and validity of the success measures for transfers. Reliable measures are stable over time. Valid measures of success for transfers should in general have correlations with institutional characteristics that are similar to those institutional characteristics found to be related to the success of beginning freshmen. At the same time since there are well understood differences between transfers and those who begin at an institution, there should be differences between the success measures of the two types of students. This validity is often considered to be construct validity as it looks at both convergent and divergent relationships.

Several cautions are appropriate when interpreting these correlations. The first caution involves the sample of institutions used in this analysis. As previously mentioned these institutions are those participants of CSRDE that voluntarily provided data concerning their transfers. In a similar manner the transfers are a select group of those who had completed at least one year of college credit at a community college. These two limitations produce a rather homogeneous sample of both institutions and transfers. They also represent a type of selection bias limiting the ability to generalize results to a broader group of institutions and transfers in general.

The second caution is the previously mentioned fact that correlations between variables based on institutional averages and proportions can be substantially different from correlations developed between variables where the unit of analysis is at the individual student level. This is known as the ecology fallacy and is discussed in many textbooks and discussions on correlations (See Appendix A-2).

The final caution is the traditional reminder that correlation does not imply causality. With the issues of confounding variables and the numerous factors operating in the context and culture of institutions, these correlations only establish suggestions of relationships that might be investigated to determine causality.

Having shared these cautions, correlations in Table 2 are presented in the spirit of discovery and to generate discussion concerning the types of institutions, and possibly the types of community college transfer students, that are more likely to be successful. If these correlations accomplish those two purposes then the efforts made by participating institutions and by CSRDE have value to how we might look at student outcomes and support benchmarking institutions across state boundaries.

Correlations between Outcomes

One of the indicators of reliability is the stability of a statistic. In other words, does a measure after one year relate closely to the same measure after two years? The following looks at that question for the measures of Success and Graduation. As shown in Tables 3 through 7, Student Success is stable from the first year and Graduation less so. These four tables show the inter-correlations within the success measures and within the graduation measures. The top diagonal of the matrix are the graduation correlations and the bottom half of the matrices are the success correlations. For Asians, Blacks, and Hispanics, correlations are reported for those institutions that averaged more than two students in the cohorts included in the analysis. For example, for the Asian cohorts, the correlation of success after one year with success after two years was .909 as shown in the bottom diagonal; the correlation of graduation after one year with graduation after two years was .336 as shown in the top diagonal.

The results in Tables 3 through 7 show that the graduation rates after three years become fairly good predictors of the graduation rates for five and six years. For example, the lowest correlation of three year graduation with five year graduation was .79 and the lowest correlation of three year graduation with six year graduation was .71. Both of these were for the Hispanic cohort for institutions with more than an average of two per year in their cohort.

The results also show that sequential year success rates tend to be quite strong and in the neighborhood of .9 or higher. The reader is reminded that this result is after removing those programs where the average cohort was reported at fewer than two students a year (a total of fewer than 14 students in the seven cohorts).

Table 3: Correlations for Asian Cohorts for Institutions Reporting More Than 2 Per Cohort⁹

		Graduation					
		1 Yr	2 Yrs	3 Yrs	4 Yrs	5 Yrs	6 Yrs
Success	1 Yr	1	0.336	0.178	-0.006	-0.042	0.020
	2 Yrs	0.909	1	0.785	0.596	0.433	0.428
	3 Yrs	0.771	0.866	1	0.881	0.793	0.712
	4 Yrs	0.751	0.831	0.917	1	0.924	0.839
	5 Yrs	0.590	0.697	0.824	0.918	1	0.895
	6 Yrs	0.480	0.579	0.711	0.800	0.848	1

Table 4: Correlations for Black Cohorts for Institutions Reporting More Than 2 Per Cohort

		Graduation					
		1 Yr	2 Yrs	3 Yrs	4 Yrs	5 Yrs	6 Yrs
Success	1 Yr	1	0.420	0.227	0.175	0.130	0.071
	2 Yrs	0.891	1	0.843	0.740	0.682	0.632
	3 Yrs	0.859	0.924	1	0.932	0.876	0.786
	4 Yrs	0.827	0.886	0.915	1	0.953	0.856
	5 Yrs	0.783	0.840	0.857	0.949	1	0.916
	6 Yrs	0.674	0.753	0.773	0.853	0.896	1

⁹ If one wanted to know the “reliability of averaging two adjacent Success ratios, the Spearman-Brown Prophecy formula would give $2*r/(1+r)$ or about .95 or higher for most of the adjacent ratios.

Table 5: Correlations for Hispanic Cohorts for Institutions Reporting More Than 2 Per Cohort

		Graduation					
		1 Yr	2 Yrs	3 Yrs	4 Yrs	5 Yrs	6 Yrs
Success	1 Yr	1	0.179	0.133	0.103	0.080	0.032
	2 Yrs	0.892	1	0.905	0.791	0.670	0.646
	3 Yrs	0.773	0.874	1	0.929	0.843	0.791
	4 Yrs	0.784	0.825	0.900	1	0.944	0.881
	5 Yrs	0.714	0.783	0.894	0.923	1	0.941
	6 Yrs	0.729	0.759	0.851	0.879	0.929	1

Table 6: Correlations for White Cohorts for Institutions Reporting More Than 2 Per Cohort

		Graduation					
		1 Yr	2 Yrs	3 Yrs	4 Yrs	5 Yrs	6 Yrs
Success	1 Yr	1	0.344	0.169	0.087	0.108	0.089
	2 Yrs	0.957	1	0.876	0.739	0.680	0.668
	3 Yrs	0.913	0.964	1	0.940	0.885	0.867
	4 Yrs	0.905	0.943	0.972	1	0.979	0.959
	5 Yrs	0.891	0.936	0.964	0.979	1	0.981
	6 Yrs	0.860	0.907	0.933	0.956	0.975	1

Table 7: Correlations for Total Cohorts

		Graduation					
		1 Yr	2 Yrs	3 Yrs	4 Yrs	5 Yrs	6 Yrs
1 Yr		1	0.402	0.219	0.127	0.122	0.102
2 Yrs		0.960	1	0.871	0.725	0.676	0.638
Success	3 Yrs	0.913	0.962	1	0.939	0.904	0.866
	4 Yrs	0.881	0.922	0.971	1	0.982	0.957
	5 Yrs	0.868	0.906	0.967	0.982	1	0.985
	6 Yrs	0.835	0.866	0.936	0.970	0.973	1

Predictive Validity

One of the issues in using student outcomes is that while the intent is on increasing the graduation rate, this student outcome measure takes a longer time to occur than many key stakeholders are inclined to wait. The desire is to look at a shorter time criteria and use this to estimate the longer time criteria. This is a measure of the statistical validity of a measure. In terms of this research the question is the ability to predict the graduation in five or six years from the success in a shorter term of time. This is shown in Tables 8 through 12 as the correlations between success in one to six years and graduation in one to six years. Where the earlier tables 3 to 7 showed the reliability (stability) as related to success and graduation, these tables show the predictive validity of success for various graduation statistics. For example, for institutions that have reported more than an average of two Asians in their transfer cohort, the average Year 1 success of the cohort is correlated .579 with its five year graduation rate and .605 with its six year graduation rate. For students reported in the Black cohort these correlations were .784 and .677.

**Table 8: Correlations for Asian Cohorts for Institutions Reporting More Than 2 Per Cohort
(n = 62, Average Cohort = 104)**

		Graduation					
		1 Yr	2 Yrs	3 Yrs	4 Yrs	5 Yrs	6 Yrs
	1 Yr	-0.070	0.332	0.535	0.688	0.579	0.605
	2 Yrs	-0.085	0.280	0.568	0.751	0.688	0.677
Success	3 Yrs	0.047	0.348	0.698	0.845	0.828	0.787
	4 Yrs	-0.020	0.429	0.726	0.899	0.896	0.847
	5 Yrs	-0.077	0.382	0.721	0.873	0.966	0.886
	6 Yrs	-0.001	0.389	0.631	0.752	0.831	0.959

**Table 9: Correlations for Black Cohorts for Institutions Reporting More Than 2 Per Cohort
(n = 84, Average Cohort = 75)**

		Graduation					
		1 Yr	2 Yrs	3 Yrs	4 Yrs	5 Yrs	6 Yrs
	1 Yr	0.048	0.544	0.737	0.785	0.784	0.677
	2 Yrs	0.000	0.531	0.737	0.797	0.797	0.750
Success	3 Yrs	0.052	0.591	0.804	0.834	0.838	0.773
	4 Yrs	0.113	0.614	0.816	0.921	0.929	0.855
	5 Yrs	0.119	0.592	0.801	0.903	0.951	0.900
	6 Yrs	0.065	0.607	0.746	0.823	0.893	0.982

Table 10: Correlations for Hispanic Cohorts for Institutions Reporting More Than 2 Per Cohort (n= 73, Average Cohort = 205)

		Graduation					
		1 Yr	2 Yrs	3 Yrs	4 Yrs	5 Yrs	6 Yrs
	1 Yr	-0.082	0.554	0.707	0.799	0.751	0.747
	2 Yrs	-0.017	0.533	0.719	0.838	0.819	0.784
Success	3 Yrs	0.070	0.604	0.815	0.900	0.907	0.847
	4 Yrs	0.133	0.638	0.802	0.922	0.911	0.872
	5 Yrs	0.095	0.623	0.793	0.888	0.958	0.920
	6 Yrs	0.044	0.603	0.758	0.843	0.905	0.968

Table 11: Correlations for White Cohorts for Institutions Reporting More Than 2 Per Cohort (n = 111, Average Cohort = 442)

		Graduation					
		1 Yr	2 Yrs	3 Yrs	4 Yrs	5 Yrs	6 Yrs
	1 Yr	-0.003	0.526	0.742	0.854	0.872	0.859
	2 Yrs	0.039	0.540	0.767	0.880	0.910	0.897
Success	3 Yrs	0.075	0.629	0.833	0.914	0.940	0.927
	4 Yrs	0.077	0.641	0.843	0.946	0.963	0.949
	5 Yrs	0.114	0.653	0.846	0.948	0.983	0.968
	6 Yrs	0.091	0.657	0.844	0.940	0.970	0.988

Table 12: Correlations for Total Cohorts
(n = 117. Average Cohort = 781)

		Graduation					
		1 Yr	2 Yrs	3 Yrs	4 Yrs	5 Yrs	6 Yrs
	1 Yr	-0.029	0.533	0.745	0.855	0.859	0.849
	2 Yrs	0.018	0.537	0.759	0.883	0.885	0.876
Success	3 Yrs	0.079	0.633	0.847	0.936	0.950	0.941
	4 Yrs	0.080	0.612	0.832	0.947	0.965	0.965
	5 Yrs	0.131	0.658	0.864	0.955	0.981	0.976
	6 Yrs	0.095	0.612	0.832	0.931	0.969	0.991

Several aspects of Tables 8 to 12 should be noted. The first is that five and six year graduation rates can be predicted by earlier success rates before the total time has passed. This is particularly true for the Total cohort and also the White cohort. It is also true for the other three ethnic cohorts but to a slightly lesser degree. The next aspect is that in the first several years the graduation rate does a much poorer job of predicting the five and six year graduation rate than does the success rate. The reader is reminded that these results show the relationship between success and graduation rates. Institutions that are more successful in the first several years, and by implication, programs that are more successful the first several years, will tend to have a higher graduation rate in five or six years than do those institutions that are less successful in the first several years. These correlations are not a measure of magnitude but reflect strength of association with the earlier success rates. Explanatory models, such as those from HLM and regression analysis are required for statements of magnitude.

Correlations Within and Between Ethnic Cohorts

The correlations between and within the ethnic groups can provide a sense of the convergent and divergent validity of the data. The question is—does the data detail provide an expected pattern when used to consider the relationship of various concepts across and within ethnic group. Success rates across ethnic cohorts should be correlated for the same year of success. These correlations should be higher within cohort than across cohorts. The correlations should be higher within year than across

years. The lowest correlations should be across cohorts and across years. This concept is adapted from the Multi-Trait-Multi-Method described by Campbell and Fisk (1959).

The data to address these questions are shown in Table 13. The two measures of success and the two measures of graduation selected were thought to represent the most interesting measures of their type. It seemed a bit complicated to select all 14 of the measures.

The correlation on the top diagonal of the matrix is an average of the correlations of the components within ethnic group. In other words, the average correlation of cohort size and the % part-time within ethnic groups (e.g. sum of the {correlation for Asians plus correlation for Blacks plus correlation for Hispanics plus correlation for Whites} divided by 4). In general, the size of the cohort has a negligible relationship with other factors. On the other hand, the % part-time has a negative correlation with the four measures of student outcomes. The two measures of success are both strongly related to the two graduation rates. (These are somewhat redundant with the data provided in Table 8 but are provided for reference with the other parts of the matrix.)

The diagonal is the consistency of the measures across ethnic groups. In other words, the correlation of the number in the cohorts is .43 across the six combinations of different cohorts taken two at a time (e.g. Asian with Black, Asian with Hispanic...Hispanic with White). This shows that large programs tend to have large numbers in the ethnic groups but this is not a strong tendency. What is a strong tendency is that there tends to be a part-time characteristic of the programs and this is consistent across ethnic groups. Also, there is a success character and this tends to be consistent across ethnic groups with programs. The likelihood of graduating in five years explains about half the variance between ethnic groups with the average correlation of .75 (the average of the correlation of Asian with Black, Asian with Hispanic...Hispanic with White).

The lower diagonal is the average of correlations across the measures between the ethnic cohorts. For example, the count for one cohort is correlated with the % part-time with another cohort. As might be expected, these have a pattern somewhat similar to the correlations above the diagonal but are a bit lower. This being a bit lower is the divergent validity and indicates that the analysis that uses the cohorts based on ethnic categories adds additional information.

Table 13: Correlations for Selected Graduation and Success Rates between Ethnic Cohorts (lower diagonal matrix and diagonal) and within Ethnic Cohorts (upper diagonal matrix)

	Count	%Part-time	1 Yr Success	2 Yrs Success	Grad in_5_Yrs	Grad in_6_Yrs
Count	0.463	0.210	0.199	0.201	0.162	0.135
% Part-time	0.216	0.877	-0.375	-0.353	-0.438	-0.416
1 YR Success	0.240	-0.372	0.773	0.913	0.747	0.722
2 YR Success	0.209	-0.358	0.726	0.792	0.803	0.777
Grad in_5_Yrs	0.182	-0.442	0.673	0.723	0.746	0.933
Grad in_6_Yrs	0.165	-0.436	0.668	0.720	0.726	0.685

Correlation with Various Independent Measures

The correlations of student success were calculated with independent variables that reflected the characteristics of the institution. These variables were taken from IPEDS data, housed at the National Center for Education Statistics (NCES). As previously noted this is not an exhaustive list of possible IPEDS variables. Because of the very large number of possible measures of student outcomes, the following is focused on student success rates using the second-year success rate, which is primarily retention. It is used as a balance of being able to anticipate graduation rates and having a short-term measure for program evaluation.

The specific variables were selected based on the previous work by McLaughlin, McLaughlin, and Howard (2013) to develop a balanced scorecard of institutional characteristics. These characteristics included demographics such as control and urbanicity, undergraduate brand such as residential, selective and expensive, student demographic characteristics such as age and race/ethnicity, academic characteristics such as % Science, Technology, Engineering, and Mathematics (STEM) and faculty salary, and wealth such as endowment/FTE student. There is no intent to claim that this is an exhaustive list all characteristics and the reader is referred to current research on predicting retention of traditional students such as the studies done by Moosai (2010), Pike and Robbins (2015), Kesha (2012), and Marcos (2015). The full set of correlations is shown in the Appendix A-4. Those institutional

characteristics with an average correlation greater than .35 with the two-year success rate are shown in Table 14 and are sorted based on their average correlation with the five cohorts.

Table 14: Correlations of Institutional Characteristics with 2-year Success Rates for Four Ethnic Cohorts and the Total Cohort where the Average Correlation Is Greater than .35 (p<.05 in bold)¹⁰

Institutional Characteristic	Cohort				
	Asian	Black	Hispanic	White	Total
% 1 yr Retention (IPEDS)	0.596	0.523	0.546	0.644	0.647
6 yr Grad Rate (IPEDS)	0.560	0.482	0.393	0.613	0.629
UGFT Asian	0.504	0.479	0.450	0.479	0.468
Average Ranked Faculty Salary	0.569	0.400	0.320	0.547	0.541
% Asian UGFT	0.471	0.424	0.390	0.453	0.441
% UG Part-time	-0.388	-0.389	-0.297	-0.538	-0.534
Size (FTE)	0.473	0.449	0.396	0.385	0.399
% Alien UGPT	0.470	0.362	0.350	0.436	0.438
Freshman Applications/UG	0.277	0.430	0.380	0.459	0.483
% UG > 24 years old	-0.430	-0.345	-0.295	-0.465	-0.459
Acceptance Rate	-0.461	-0.481	-0.428	-0.292	-0.313
UGFT Hispanic	0.258	0.371	0.400	0.411	0.394
% White UGPT	-0.338	-0.405	-0.368	-0.352	-0.316

The most striking aspect of Table 14 is that the best predictors of transfer student success is the success of the IPEDS cohort—the first-time, full-time, degree-seeking student. It is also evident that there is a relationship of student success to the measures of a traditional selective and larger university. These factors are less strong for transfers, however, than they are for traditional freshmen. One of the strongest predictors of freshman success has traditionally been the endowment per FTE UG student and while this

¹⁰ The cut-point of .35 was used since this is about 10% explained variance and there was a break at that point and finally this seemed to include enough variables to indicate how the institutional characteristics related to the 2-yr success rate. The full data are in the Appendix.

did have a positive relationship with four of the five cohorts, its average correlation was below .35, as shown in the Appendix. Table 15 shows the institutional characteristics that had the stronger correlations with the five-year graduation rates.

Table 15: Correlations of Institutional Characteristics with 5-year Graduation Rates for Four Ethnic Cohorts and the Total Cohort where the Average Correlation Is Greater than .35 (p<.05 in bold)¹¹

	Asian	Black	Hispanic	White	Total
6 yr Grad Rate	0.589	0.575	0.505	0.627	0.681
% Retention	0.532	0.524	0.526	0.565	0.605
% Alien UGPT	0.509	0.502	0.432	0.485	0.514
UGFT Asian	0.460	0.535	0.462	0.453	0.442
Average Ranked Faculty Salary	0.429	0.457	0.446	0.465	0.518
% UG Part-time	-0.330	-0.442	-0.405	-0.599	-0.538
% UG > 24	-0.410	-0.423	-0.394	-0.531	-0.501
Size (FTE)	0.495	0.531	0.410	0.362	0.412
apps/UG	0.295	0.469	0.440	0.522	0.469
% Asian UGFT	0.382	0.438	0.397	0.385	0.373
ACT	0.455	0.411	0.308	0.269	0.361
Acceptance Rate	-0.446	-0.401	-0.384	-0.311	-0.241
\$ Research & Svc /TT Faculty	0.499	0.445	0.257	0.242	0.286
UGPT_Alien	0.365	0.425	0.304	0.300	0.333
% FTFTDS Cohort	0.407	0.306	0.266	0.400	0.330
Tuition/state Dependency	-0.518	-0.406	-0.310	-0.260	-0.197
Tuition	0.302	0.265	0.301	0.320	0.352
% White UGPT	-0.298	-0.344	-0.355	-0.339	-0.175

These institutional characteristics again reinforce the fact that institutions that do well with transfers are also those that do well with the traditional college student. Factors related to higher graduation rates include higher ACT scores, lower acceptance rates, larger proportions of traditional undergraduates, and

¹¹ The cut-point of .35 was used since this is about 10% and there was a break at that point and finally this seemed to include enough variables to indicate how the institutional characteristics related to the 2-yr success rate. The full data are in the Appendix.

higher tuition. Again however, these factors are not as strongly related to graduation of transfers as they have been found to be related to the graduation of the more traditional cohort in previous studies.

Discussion of Correlations

In terms of the correlations between measures for transfer success and transfer graduation, it is apparent that these measures become extremely stable within a limited number of years. For graduation after two to three years, correlations approach or exceed .9. For measures of transfer student success correlations, particularly between adjacent years, exceed .9 from the first year. As one would expect correlations go down slightly as the number of years between variables increases.

It is not the intent of this research to draw conclusions concerning institutional characteristics of those institutions that are successful in supporting the success of their community college transfers. The preceding correlations are very encouraging with regards to the validity of the data provided by the institutions that provided transfer data to CSRDE. Many of the correlations are highly statistically significant. In some ways institutions that are successful in retaining and graduating beginning students are also successful in retaining and graduating transfers. This occurs, for example, with measures that are identified as part of the institutional undergraduate brand. In other cases, the variables are much more highly related to freshmen outcomes than to those outcomes for transfer students. This is the case, for example, with variables related to institutional financial characteristics.

Methodology 3: Hierarchical Linear Modeling

HLM is appropriate for this study because student groups are nested within institutions, violating the assumption of independence associated with general linear models. Building a model at the institutional-level can mask individual differences, an effect called “ecological fallacy” (Hu & Kuh, 2003; Kreft & deLeeuw, 1998), while attaching institution-level characteristics to an individual or group violates the OLS assumption that observations are independent of one another (Ethington, 1997). Since this study examines the impact of institutions on the academic success of student groups, HLM was used to analyze the proportion of variance in two-year success rates explained within institutions across racial groups (level-1) and between institutions (level-2). The null, or “unconditional,” model contained no independent variables [$\text{TwoYS}_{ij} = \beta_{0j} + e_{ij}$ & $\beta_{0j} = \gamma_{00} + \mu_{0j}$]. The random effects model contained level-1 variables race and percent transfer cohort as part-time [$\text{TwoYS}_{ij} = \beta_{0j} + \beta_{1j}\text{Ethnicity}_{ij} + \beta_{2j}\text{Part_Time}_{ij}$

Measuring the Success of Transfer Students

+ e_{ij} & $\beta_{0j} = \gamma_{00} + \mu_{0j}$ & $\beta_{1j} = \gamma_{10}$ & $\beta_{2j} = \gamma_{20}$]. The final model added level-2 variables as fixed effects 6-year graduation rate, acceptance rate, and percent faculty as full-time [$\text{TwoYS}_{ij} = \beta_{0j} + \beta_{1j}\text{Ethnicity}_{ij} + \beta_{2j}\text{Part_Time}_{ij} + \beta_{3j}\text{Acceptance}_{ij} + \beta_{4j}\text{6YR}_{ij} + \beta_{5j}\text{Faculty}_{ij} + e_{ij}$ & $\beta_{0j} = \gamma_{00} + \mu_{0j}$ & $\beta_{1j} = \gamma_{10}$ & $\beta_{2j} = \gamma_{20}$ & $\beta_{3j} = \gamma_{30}$ & $\beta_{4j} = \gamma_{40}$ & $\beta_{5j} = \gamma_{50}$]. These variables were selected since their condition index was below 30, as set by Belsley, Kuh, and Welsch (2005), indicating that strong collinearity was not a problem in the model. Adding additional variables to the model violated this collinearity standard and excessive intercorrelation of independent variables can increase the standard error of the estimates and make parameter estimates inefficient.

Although additional models are available in multilevel analysis (i.e., intercepts-and-slopes-as-outcomes), the intercepts-as-outcomes model was used since it is the most parsimonious manner to remove the assumption of uncorrelated error terms. Each model contained the same dependent variable, two-year success rate. The between- and within-groups effects were used to compute the intraclass correlation (ICC). Pseudo R^2 was calculated according to Snijders and Bosker's (1994) metric: $(\sigma^2_{\text{null}} - \sigma^2_{\text{model}})/\sigma^2_{\text{null}}$. All statistics were conducted in SPSS, version 23 and statistical significance was established at $\alpha = .05$. Continuous data are represented as mean (standard deviation). Success in two years was selected based on the previous analyses and consideration of institutional perspectives.

Results – HLM

There were 330 racial groups nested within 113 institutions, with 2.92(1.16) groups per institution. The average two year success rate for each racial group was 0.76(.07) for Asian ($n = 62$), 0.68(.09) for Black ($n = 84$), 0.74(.07) for Hispanic ($n = 73$), and 0.74 (.07) for White ($n = 111$). Descriptive statistics for all variables entered into the model can be seen in Table 16.

Table 16. Descriptive Statistics for HLM Variables at Level-1 and Level-2

	N	Minimum	Maximum	Mean	Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Statistic
@2YRSuccess	330	.38	.95	.7267	.08206
<i>Level-1</i>					
Transfer cohort as Part-Time	330	.00	.49	.2091	.11529
<i>Level-2</i>					
Acceptance Rate	330	.20	1.00	.6540	.15570
6 yr Grad Rate	330	.24	.93	.5471	.15430
% Faculty as FT	330	.66	1.00	.8607	.09857

Sample (n = 330) consists of Asian (n = 62), Black (n = 84), Hispanic (n = 73), and White (n = 111) groups nested within institutions (n = 113).

The null model assessed the degree of between-group variance in the two-year success rate. In this study, the null model had an ICC of .66, which means that 66% of the model variance in two-year success rates for groups of transfer students was explained by institutional characteristics. It can also be said that two-year success rates clustered by institution, meaning that two transfer groups randomly selected from the same school were more likely to have similar rates than a pair of randomly selected groups representing different schools.

The random effects model fully considered all level-1 variables while controlling level-2 variables. In this model, 73.58% of the variation in two-year success rate was accounted for by the institution and the race/ethnicity ($p < .001$) and part-time percentage of the group ($p < .001$). A pseudo R^2 of .39 indicated that these level-1 predictors accounted for 39% of the explainable variance at level-1. In the final model, acceptance rate ($p < .001$), six-year graduation rate ($p < .001$), and percent faculty as full-time ($p < .05$) further contributed to the variance in two-year success rates and help explain differences between two-year success rates. Sixty-three percent of the variation in two-year success rate was accounted for by the final model (i.e., the institution, the race/ethnicity and part-time percentage of the group, and the acceptance rate, six-year graduation rate, and percent faculty as full-time). By adding institution-level predictors that accounted for variation among the institutions, less variation existed in the random

intercept for this model and resulted in a lower ICC. A pseudo R^2 of .50 indicated that the addition of level-2 variables to the model accounted for 50% of the between-group variance in two-year success rates. Deviance, as indicated by 2LL model, decreased for each model, indicating improved model fit. These results are show in Table 17.

Table 17. Group-Level and Institution-Level Predictors of Two Year Success Rates for Transfer Cohorts

	Null		Level-1		Levels-1 & -2	
	Null	St. Error	Random Intercepts	St. Error	Intercepts-as-Outcomes	St. Error
<i>Fixed Effects</i>						
Intercept	.72**	.01	.77**	.01	.84**	.05
Asian ¹			.01	.01	-.01	.01
Black ¹			-.06**	.01	-.06**	.01
Hispanic ¹			-.01	.01	-.01	.01
Transfer Cohort as Part-Time			-.18**	.04	-.10*	.04
Acceptance Rate					-.12**	.03
6 yr Grad Rate					.25**	.04
% Faculty as FT					-.16*	.06
<i>Deviance</i>	-855.76		-961.69		-1002.73	
<i>Covariance Parameters</i>						
Residual	.0023**		.0014**	<.01	.0013**	<.01
Pseudo R^2				.39		.43
Intercept	.0046**		.0039**	<.01	.0023**	<.01
Pseudo R^2						.50

¹Referent group is white; * p<.05; **p<.001

Discussion of HLM

The findings from this analysis hold statistical and practical significance, contributing to our understanding of transfer success rates. First, the data are sufficiently reliable and valid to support HLM, and HLM is an appropriate model for looking at the role of institutional characteristics in explaining success statistics for student groups. Second, these models extend our current understanding of the

relationship between institutional characteristics and academic outcomes and provide an important step towards (a) understanding transfer success in a hierarchical framework and (b) analyzing success of transfer groups by race/ethnicity.

Areas for Further Study

Based on the results of the analysis presented above, it is evident that cohorts of transfers and success rates are in fact related to institutional characteristics and average student characteristics. In this sense, they are similar to the traditional IPEDS cohort. These analyses also show, however, that the outcomes for transfer cohorts are different from the outcomes for the IPEDS cohort. These analyses are strongly encouraging to those institutions who want to benchmark themselves on the success of their transfer students. The findings should also be strongly encouraging to those who want to do additional analyses on similar data. This leads to topics for areas of further study. The following are some ideas that have been generated in discussing these data and results with colleagues.

1. Further research can be done with extending the HLM analyses. In projecting proportions it is frequently desirable to use logistic regression, particularly when some of the institutions are operating close to the extremes of zero and one. Also, we used one set of assumptions in the HLM. There are alternative assumptions that would be proscribed by a researcher's hypotheses. Finally, it is likely that some would prefer to model other criteria.
2. As shown in the first analysis, data are available for the sequence of years from one year to seven years. It is very possible that different institutional characteristics are associated with student success at different points in their enrollment. This suggests the use of time – outcome methodologies such as survival analysis.
3. While there currently is a limited sample for these data, it is still feasible for institutions to form comparator groups based on their characteristics. This is particularly feasible for the medium to larger public institutions as these tend to dominate these data. It is also hoped that the statistical results of these analyses will encourage other institutions and groups of institutions to provide their data to CSRDE so that benchmarking can produce increasingly stable results.
4. A major area of further study would be to investigate other possible independent variables. One of the unused aspects of the CSRDE data is the identification of primary feeder community colleges for these institutions. A likely extension of the independent dataset would be

to obtain the IPEDS data for these feeder community colleges and use those data to form independent variables for use in a regression or HLM model. With the appropriate approval of the institutions these data could also be extended to include the demographic aspects of the context of the institutions.

5. A final suggested area of research is to make adjustments to, and/or create additional cohorts of transfers students. One cohort would be students transferring from four-year institutions. Another cohort adjustment might be to take into consideration whether students had completed an associate degree or not at the community college. Another cohort adjustment might be to create cohorts based on academic discipline. This represents a deconstruction of the previously mentioned Student Achievement Measure.

Summary

This study has investigated the likely value of data on transfer students from community colleges to four-year institutions who had completed the equivalent of at least one academic year prior to their transfer. This investigation was in three parts: the computation of means, the computation of correlations, and the development of HLM models. These three methodologies tend to be the majority of those found in the analysis of institutional data on student success. The reader is reminded that the sample of institutions is restricted to institutions that are participating in the CSRDE data exchange and furthermore provided sufficient data about their transfer students to be included in the study.

The data from the institutions includes student persistence and graduation rates for the selected cohorts for seven years after their enrollment to include enrollment in the fall following the seventh year. Furthermore these data are by Total, by Race/Ethnicity and by Gender. Data were available for seven cohorts for the seven years of beginning transfer students. Data were used after data were averaged over the cohorts provided by the institutions. This was done to use the data in their most stable form because of the modest sample size of 117 institutions and because of the initial exploratory nature of this research. The means by year indicated that the majority of the graduations occurred in the first five years after enrollment. They also indicated, however, that some students did continue to persist and enroll, most likely not continuously, and graduate throughout the time on which data were available.

The correlations reinforced some of the traditional knowledge about student success, i.e., those institutions that have larger portions of transfers and that have larger proportions of their enrollments as part-time also tend to have the lower success rates in terms of retention and graduation. The correlations also showed that these data have convergent validity in that relatively high initial success rates for transfers are associated with higher success rates after six or seven years. Relatively high success rates for transfers are also significantly related to success rates for first-time full-time degree-seeking students. However, there are differences in the correlations. Success rates for beginning student cohorts are related to more institutional characteristics than are the success rates for transfer students. One possible explanation for this higher association is that beginning freshmen are more likely to be mobile and select institutions based on institutional brand, while transfer students are traditionally thought to be more geographically restricted in their options.

The presence of similarities and differences also occurred in the HLM models. First, and most importantly, the student and institutional characteristics had a strong to very strong ability to anticipate student success and outcomes by institution and by ethnicity within institution. Rather than build interpretations of what these relationships mean, this research points to the strength of those relationships as the likely value of these data in building a better understanding of success outcomes for transfer students.

Viewing these three sets of results as an overall consideration, these data, and similar data built on focus descriptions of transfer students would seem to be a very valuable step for the institution that wants to describe the outcomes of transfer students within their programs. These data would also seem to be a very valuable next step for researchers who want to discuss benchmarks for institutional outcomes and who want to generate hypotheses for looking at programmatic success relative to transfers from community colleges.

References

- Belsley, D. A., Kuh, E., & Welsch, R. E. (2005). *Regression diagnostics: Identifying influential data and sources of collinearity* (Vol. 571). John Wiley & Sons.
- Campbell, D.T., & Fiske D.W. (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.

- Ecological fallacy. (n.d.) Definition. Wikipedia. Retrieved from https://en.wikipedia.org/wiki/Ecological_fallacy
- Ethington, C. A. (1997). A hierarchical linear modeling approach to studying college effects. *Higher Education Handbook of Theory and Research*, (12), 165-194.
- Freedman, D. A., (1999, Oct). *Ecological inference and the ecological fallacy*, International Encyclopedia of the Social & Behavioral Sciences, Technical Report No. 549. Retrieved from <http://web.stanford.edu/class/ed260/freedman549.pdf>
- Hu, S., & Kuh, G. D. (2003). Maximizing what students get out of college: Testing a learning productivity model. *Journal of College Student Development*, (44), 185- 203.
- Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. London: Sage.
- Lee, K. (2012). *An analysis of the institutional factors that influence retention and 6-year graduation rates at historically black colleges and universities* (thesis). Sanford School of Public Policy, Duke University, Durham, NC.
- McLaughlin, G.W., McLaughlin, J.S. & Howard, R.D. (2013). Reference group formation using the nearest neighbor method. *Planning for Higher Education*, 41(2), 136-158.
- McLaughlin, G.W., McLaughlin, J.E., and McLaughlin, J.S. (2013). Diversity indices: Research on metrics for measuring diversity in US higher education. In N. Huber, P. Pohlenz, & R. Krempkow (Eds.), *Diversity in Science* (pp. 185-204). Berlin: UVW-Verlag.
- Moosai, S. (2010). A prediction model for community colleges using graduation rate as the performance indicator (dissertation). College of Education, Florida Atlantic University, Boca Raton, FL.
- Pike, G. R. & Robbins, K. (2015, May). Using panel data to identify the effects of expenditures on graduation rates. Presented at forum of Association for Institutional Research, Denver, CO.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*, Second edition. Thousand Oaks, CA: Sage Publications.
- Snijders, T. A. B., & Boskers, R. J. (1994). Modeled variance in two-level models. *Sociological Methods and Research*, (22), 342-363.
- Student Achievement Measure. (n.d.). *SAM bachelor's degree-seeking model methodology*. Retrieved from www.studentachievementmeasure.org/resources .
- Thorndike, E. L., (1939, January). *On the fallacy of imputing the correlations found for groups to the individuals or smaller groups composing them*, *The American Journal of Psychology*, 52 (1): 122-124. Retrieved from http://www.jstor.org/stable/1416673?seq=1#page_scan_tab_contents

Measuring the Success of Transfer Students

Trochim, W. M.K. (2006). *Two research fallacies*. Research Methods: Knowledge Base. Retrieved 4/15/15 from <http://www.socialresearchmethods.net/kb/fallacy.php>

Velazquez, M. (2015). The relationship between institutional characteristics and first-year retention rates. Presented at forum of Association for Institutional Research, Denver, CO.

Appendix

Appendix A-1: The CSRDE Transfer Data Exchange

About the CSRDE Transfer Survey

CSRDE members expressed the need in 2005 to evaluate the success of transfer students within their institutions, but developing these data was a challenge due to the variation in the characteristics of transfer students. Some came from two-year institutions, others transferred from four-year institutions, and other students had simply taken several courses at a two-year college while waiting to enroll in a four-year institution. Some of these students intended to earn a degree while others were merely interested in taking an occasional course or two. Discussions with members of the Consortium revealed that the students of primary concern were those transferring from two-year institutions who had taken at least the equivalent of one academic year of courses at the two-year institution prior to the transfer.

A data sharing protocol was developed with the cohort defined as the group of degree-seeking students entering the senior institution for the first time in the fall term who had completed at least 30 hours at a community college—the equivalent of at least one year of study. This was seen as a measure of those who were attending the community college as a substantial step in their higher education pursuit and eliminated those who only took a course or two at a community college.

The CSRDE Community College Transfer survey is an optional data collection in which Consortium members may participate at no cost. The survey tracks year-to-year retention and graduation data on ten cohorts of student for ten years, and includes first-year graduation through 11th year retention data. The data includes sub-cohorts by gender and race/ethnicity. The survey upon which this paper is based includes the 2003-2012 cohorts; however, the analyses for this research comprise only the 2006-2012 cohorts. Both full-time and part-time students make up the cohorts, and the survey also collects the percentage of each cohort who enrolled as part-time in their first semester.

In addition, participating institutions also submit data related to the entire transfer population at their institution, such as number of students who applied, were admitted and enrolled for a particular cohort; minimum GPA required; minimum grade accepted; maximum number of credits that may be transferred; and the top five feeder schools. Some of these data are taken from Part D of the Common Data Set (CDS) for those institutions participating in that survey, or they are provided in the CSRDE transfer survey template. A brief description of the CDS initiative may be found at <http://www.commondataset.org>.

About the CSRDE

The Consortium for Student Retention Data Exchange (CSRDE) at the University of Oklahoma is an association of two-year and four-year institutions with the common goal of achieving the highest possible levels of student success through collaboratively sharing data, knowledge and innovation. Founded in 1994 by a small group of Institutional Research directors as a data exchange of college retention and graduation data, the first CSRDE report was published in May of 1995.

Membership in CSRDE is open to two-year and four-year, public and private nonprofit institutions. Participants provide retention and graduation data on first-time students to be included in the annual CSRDE Retention Reports and made available exclusively to member institutions. All surveys include sub-cohorts of data by gender and race/ethnicity. Access to peer data for benchmarking is available online through the Quick Query System.

Measuring the Success of Transfer Students

In addition to the survey for first-year students and the Transfer survey upon which this paper is based, four-year institutions may also participate in the Science, Technology, Engineering, and Mathematics (STEM) survey at no cost. This survey was developed in the late 1990's and originally funded through a National Science Foundation grant.

The retention survey for two-year institutions includes full-time and part-time degree/certificate-seeking cohorts—reported separately—and also tracks retention and graduation data on students taking remedial courses. As in the survey for four-year institutions, ten years of cohorts are followed for ten years. More information about the CSRDE surveys and reports is found on the website at <http://csrde.ou.edu>.

In addition to publishing four annual retention reports, the Consortium also hosts the National Symposium on Student Retention each fall and a monthly webinar series. Members have access to the Knowledgebase where all peer-reviewed papers presented at previous conferences have been archived. In 2013, the CSRDE published *Building Bridges for Student Success: A Sourcebook for Colleges and Universities*, designed as a dynamic electronic document providing higher education practitioners and researchers with an expanding resource of practices and factors that influence student success. Colleagues from member institutions receive a discount on the purchase of the Sourcebook.

For more information about the Consortium, please visit our website at <http://csrde.ou.edu> or email csrde@ou.edu.

Appendix A-2: The Ecological Fallacy

“An ecological fallacy is a [logical fallacy](#) in the interpretation of [statistical data](#) where [inferences](#) about the nature of individuals are deduced from inference for the group to which those individuals belong. (https://en.wikipedia.org/wiki/Ecological_fallacy)

An ecological fallacy is not a statistical or methodological issue, but an issue of misinterpretation. This happens when the results of analyses using group averages or summary data are inferred to the individuals in the groups under study in the original analysis. The inferences can reflect supposed attributes of the individuals in the groups or the strength of relationships between two variables (correlation) or in regression analyses where a group of variables are used to predict a characteristic or outcome. Below, examples are given for each of these types of ecological fallacies.

Trochim (2006, in: <http://www.socialresearchmethods.net/kb/fallacy.php>) provides an example:

For instance, assume that you measured the math scores of a particular classroom and found that they had the highest average score in the district. Later (probably at the mall) you run into one of the kids from that class and you think to yourself "she must be a math whiz." Aha! Fallacy! Just because she comes from the class with the highest average doesn't mean that she is automatically a high-scorer in math. She could be the lowest math scorer in a class that otherwise consists of math geniuses!

So, when an institution is identified by *U.S. News and World Report* as the best in the country—or any ranking using institutional level data (i.e., IPEDS data)—one would be committing an ecological fallacy if he/her were to conclude that all students attending that institution were among the best in the U.S.

In a perfect world, all students would be retained and graduated at all institutions. If this were the case, there would no variance in these output measures and hence there would be no relationship (r or $R = 0$) between institutional characteristics or entering student characteristics and these output measures. However, we don't live in a perfect world and graduation and retention rates vary by institution. As pointed out by Thorndike (1939), correlations developed using institutional or group data will usually be much stronger than those found in the same analysis using unit record data, with the relationships often approaching 0. In this discussion, Thorndike's primary concern is the inappropriate inference that the relationship found using group data was what would be found at the individual level. In a footnote, he concludes:

If the fallacy was encouraged only by incompetent scientists, a note about it in this JOURNAL would not be needed. (p. 122).

Freedman (1999) refers to this same issue as *ecological correlations* and then expands the discussion to *ecological regression*. As he points out, regression analyses using group data will result in inflated R 's consistent with the level of inconsistency between the groups on the predictor variables. Freedman also provides several regression techniques that can reduce the discrepancy. His conclusions, some sixty years later, mirror that of Thorndike's:

Aggregate data are often easier to obtain than data on individuals, and may offer valuable clues about individual behavior. Ecological inferences will therefore continue to be made. The problems of confounding and aggregation bias, however, are unlikely to be resolved in the proximate future. (p. 5)

What does this mean for the interpretation of analyses conducted on higher education data? Both types of analysis are valid; it is the interpretation of the results that can be fallacious. If the concern is at the institutional

Measuring the Success of Transfer Students

level, then using institutional averages in the correlation or regression equations is appropriate. However, if the focus of the question is the individual student, then the analysis should be conducted using student unit record data. If only summary or group data are available, key to making any kind of inference to the individuals in the group (students or institutions) is your confidence in the assumption that there is very little if any variance within or between the groups as related the criterion measures or variables.

Appendix A-3: Comments on Modeling Student Success with Institutional Characteristics

U.S. Department of Education

(<http://www.ed.gov/college-affordability/college-ratings-and-paying-performance>)

“Your measurements will hurt colleges who specialize in serving disadvantaged students, like HBCUs. And colleges will respond by excluding disadvantaged students. Are you sacrificing low-income students for middle-income students who are more likely to succeed on your measures?”

1. It bears repeating: we need all students seeking higher education to access and graduate from college at higher rates in order for our economy and society to reap the benefits of postsecondary education.
2. Colleges that prioritize enrolling and graduating disadvantaged students have a special place in our country. That's why our performance ratings will include factors recognizing colleges for helping disadvantaged students succeed, such as for helping more Pell-eligible students enroll and graduate.
3. The Department will engage with the public in designing this system in a way that works for everyone and is reflective of the needs of the community... We are interested in ideas and suggestions from *all* stakeholders: students and families, college faculty and administrators, state education leaders, business and industry, researchers, associations, innovators, philanthropies, consumer interest groups, and other public advocates.”

The National Postsecondary Education Cooperative

(<http://nces.ed.gov/pubs2010/2010832.pdf>)

Suggestions for Improving the IPEDS Graduation Rate Survey Data Collection And Reporting, Report of The National Postsecondary Education Cooperative, 2010, p 17.

Detailed Findings

An analysis of the qualitative data yielded four themes including problems or issues with

1. Counting and defining the composition of an initial cohort of students;
2. Counting and defining who is a completer;
3. Understanding the length or time of completion; and
4. Incorporating students who transfer out of institutions.

Inside Higher Education

'Whither Thou Goest, I Will Go'

November 20, 2014

By Alexandra W. Logue

There is a large group of students — often overlooked — whose completion of college we need to better track and encourage: transfer students. We need to do a better job of collecting and following transfer students' data and of instituting policies that help them to graduate, such as ensuring that their credits transfer. There are many reasons that these students deserve our full attention.

For several years there has been a national priority, set by President Obama, to increase the percentage of people in the United States who have college degrees, both by increasing the number of people who go to college and by increasing the percentage of college students who finish. The United States is now only 16th in the world in the percentage of young adults with college degrees, and the percentage of U.S. jobs needing college degrees is growing faster than the supply. We need to make sure that each college student completes. However, one of the largest subgroups of students — transfer students — is frequently being ignored.

That transfer students constitute a huge group is indisputable. As early as 1999, Clifford Adelman, then at the U.S. Department of Education, in his now-classic report "Answers in the Tool Box," was noting an "increasing tendency, overlooked in both policy and research, for students to attend two, three, or more colleges in the course of their undergraduate careers." A 2012 study by the National Student Clearinghouse showed that approximately one-third of all students who began college at any level transferred at least once within five years. ...

Appendix A-4: Correlations of All Institutional Characteristics with 2-Yr Success Rates

(Decreasing average correlations)

Institutional Characteristics	Asian	Black	Hispanic	White	Total	Absolute Average
% Retention	0.596	0.523	0.546	0.644	0.647	0.591
6-yr Grad Rate	0.560	0.482	0.393	0.613	0.629	0.535
UGFT Asian	0.504	0.479	0.450	0.479	0.468	0.476
\$ Faculty Sal	0.569	0.400	0.320	0.547	0.541	0.475
% Asian UGFT	0.471	0.424	0.390	0.453	0.441	0.436
% UG Part-time	-0.388	-0.389	-0.297	-0.538	-0.534	0.429
Size (FTE)	0.473	0.449	0.396	0.385	0.399	0.421
% Alien UGPT	0.470	0.362	0.350	0.436	0.438	0.411
apps/UG	0.277	0.430	0.380	0.459	0.483	0.406
% UG > 24	-0.430	-0.345	-0.295	-0.465	-0.459	0.399
Acceptance Rate	-0.461	-0.481	-0.428	-0.292	-0.313	0.395
UGFT Hispanic	0.258	0.371	0.400	0.411	0.394	0.367
% White UGPT	-0.338	-0.405	-0.368	-0.352	-0.316	0.356
% student loan	-0.394	-0.299	-0.375	-0.278	-0.259	0.321
% STEM	0.460	0.327	0.222	0.286	0.308	0.321
ACT	0.462	0.309	0.240	0.261	0.298	0.314
Research & Svc \$/TT Faculty	0.478	0.359	0.187	0.235	0.266	0.305
% Asian UGPT	0.270	0.310	0.271	0.332	0.334	0.303
UGPT Alien	0.329	0.312	0.232	0.279	0.286	0.288
% FTFTDS Cohort	0.389	0.252	0.171	0.313	0.310	0.287
UGFT Other Known	0.239	0.307	0.221	0.255	0.255	0.255
\$ Endowment/FTE Student	0.362	0.227	0.152	0.244	0.269	0.251
Tuition	0.262	0.170	0.209	0.299	0.301	0.248
Tuition/state Dependency	-0.349	-0.316	-0.124	-0.176	-0.188	0.231
UGPT Asian	0.206	0.277	0.214	0.226	0.226	0.230
% Students Pell	-0.315	-0.170	-0.011	-0.264	-0.269	0.206
% Black UGPT	-0.169	-0.068	-0.161	-0.359	-0.264	0.204
UGFT White	0.244	0.214	0.134	0.185	0.215	0.198
% Students State Aid	0.245	0.229	0.328	0.065	0.105	0.194
Instruct & Aca Spt \$/FTE Student	0.247	0.177	0.038	0.232	0.243	0.188
% White UGFT	-0.097	-0.246	-0.273	-0.171	-0.134	0.184
% Transfers	-0.351	-0.176	-0.159	-0.109	-0.117	0.183

Measuring the Success of Transfer Students

UGPT_White	-0.127	-0.115	-0.136	-0.252	-0.242	0.174
% Black UGFT	-0.141	-0.052	-0.083	-0.345	-0.248	0.174
% Hispanic UGFT	-0.027	0.174	0.173	0.304	0.229	0.170
% Alien UGFT	-0.335	-0.147	-0.060	-0.144	-0.148	0.167
% Degrees as PhD	0.342	0.145	0.124	0.105	0.107	0.165
UGPT Black	-0.088	-0.102	-0.018	-0.251	-0.217	0.135
UGPT Hispanic	0.032	0.185	0.124	0.170	0.139	0.130
% Other UGPT	-0.310	-0.125	-0.125	-0.019	-0.055	0.127
Financial Viability	0.143	0.093	0.088	0.188	0.118	0.126
Student/Faculty Ratio	-0.032	0.112	0.132	0.207	0.170	0.118
Residential Capacity/FT UG	0.248	0.126	-0.060	0.094	0.111	0.104
Administrator/FTE Student	0.211	0.143	-0.008	0.075	0.092	0.103
Faculty as % Staff	-0.142	-0.106	-0.032	-0.120	-0.112	0.102
Enrollment Growth(2007-2012)	-0.246	0.021	-0.158	-0.045	-0.063	0.098
% Hispanic UGPT	-0.134	0.101	0.062	0.223	0.138	0.078
UGPT Other Known	-0.105	0.026	-0.041	-0.130	-0.123	0.075
UGFT Black	0.119	0.130	0.135	-0.028	-0.007	0.070
Institutional & StuSvcs \$/FTE Student	-0.005	-0.016	-0.135	0.256	0.217	0.063
Net Income	0.170	0.024	-0.113	0.128	0.094	0.061
% Faculty as FT	0.216	0.006	0.035	-0.014	0.011	0.051
UGFT Alien	-0.081	0.125	0.125	0.024	0.021	0.043
% Other UGFT	-0.111	-0.053	-0.028	0.006	-0.027	0.042
% FT Faculty Ranked	-0.024	-0.132	-0.128	0.098	0.081	0.021
Average Loan of FTFT w FA	-0.015	0.017	-0.143	0.027	0.039	0.015

Appendix A-5: Correlations of all institutional Characteristics with 5-Year Graduation Rates

Institutional Characteristics	Asian	Black	Hispanic	White	Total	Absolute Average
6-yr Grad Rate	0.589	0.575	0.505	0.627	0.681	0.596
% Retention	0.532	0.524	0.526	0.565	0.605	0.551
% Alien UGPT	0.509	0.502	0.432	0.485	0.514	0.488
UGFT Asian	0.460	0.535	0.462	0.453	0.442	0.470
\$ Faculty Sal	0.429	0.457	0.446	0.465	0.518	0.463
% UG Part-time	-0.330	-0.442	-0.405	-0.599	-0.538	0.463
% UG > 24	-0.410	-0.423	-0.394	-0.531	-0.501	0.452
Size (FTE)	0.495	0.531	0.410	0.362	0.412	0.442
apps/UG	0.295	0.469	0.440	0.522	0.469	0.439
% Asian UGFT	0.382	0.438	0.397	0.385	0.373	0.395
ACT	0.455	0.411	0.308	0.269	0.361	0.361
Acceptance Rate	-0.446	-0.401	-0.384	-0.311	-0.241	0.357
Expenditures Research&Svc /TT Faculty	0.499	0.445	0.257	0.242	0.286	0.346
UGPT Alien	0.365	0.425	0.304	0.300	0.333	0.345
% FTFTDS Cohort	0.407	0.306	0.266	0.400	0.330	0.342
Tuition/state Dependency	-0.518	-0.406	-0.310	-0.260	-0.197	0.338
Tuition	0.302	0.265	0.301	0.320	0.352	0.308
% White UGPT	-0.298	-0.344	-0.355	-0.339	-0.175	0.302
UGFT Hispanic	0.135	0.342	0.332	0.339	0.316	0.293
\$ Endowment/FTE Student	0.302	0.308	0.289	0.234	0.268	0.280
UGFT Other Known	0.210	0.372	0.294	0.207	0.257	0.268
Expend Instruction & Aca Spt/FTE Student	0.342	0.265	0.206	0.237	0.281	0.266
%STEM U	0.375	0.306	0.213	0.233	0.202	0.266
% Student Loan	-0.284	-0.270	-0.346	-0.174	-0.230	0.261
UGFT White	0.325	0.317	0.171	0.193	0.287	0.259
% Students Pell	-0.257	-0.281	-0.101	-0.237	-0.362	0.248
% Black UGPT	-0.040	-0.194	-0.279	-0.269	-0.407	0.238
% Transfers	-0.389	-0.256	-0.231	-0.178	-0.135	0.238
% Degrees as PhD	0.391	0.222	0.244	0.122	0.159	0.228
% Alien UGFT	-0.319	-0.264	-0.163	-0.193	-0.179	0.224
Faculty as % Staff	-0.290	-0.222	-0.214	-0.154	-0.110	0.198
% Faculty as FT	0.311	0.161	0.190	0.146	0.143	0.190
% Black UGFT	0.013	-0.154	-0.191	-0.230	-0.379	0.188

Measuring the Success of Transfer Students

UGPT_White	-0.084	-0.103	-0.181	-0.324	-0.230	0.184
% Other UGPT	-0.323	-0.164	-0.078	-0.176	-0.178	0.184
% Asian UGPT	0.070	0.224	0.187	0.204	0.207	0.179
Administrator/FTE Student	0.316	0.224	0.124	0.118	0.103	0.177
Residential Capacity/FT UG	0.338	0.182	0.086	0.203	0.059	0.174
% Students State Aid	0.165	0.120	0.233	0.105	0.118	0.148
UGPT Black	0.043	-0.116	-0.108	-0.269	-0.242	0.138
FTE UG Enrollment Growth (2007-2012)	-0.208	-0.040	-0.147	-0.095	-0.176	0.133
Net Income	0.253	0.100	0.037	0.133	0.138	0.132
UGPT Asian	0.056	0.214	0.129	0.122	0.132	0.131
UGPT Other Known	-0.167	-0.006	-0.077	-0.226	-0.157	0.127
\$ institutional & StuSvcs \$/FTE Student	0.024	0.057	0.067	0.206	0.125	0.096
% White UGFT	-0.037	-0.125	-0.217	-0.136	0.042	0.094
% Hispanic UGFT	-0.149	0.099	0.135	0.204	0.144	0.086
Student/Faculty Ratio	-0.062	0.084	0.134	0.155	0.111	0.084
Financial Viability	0.110	0.015	0.066	0.113	0.077	0.076
UGPT Hispanic	-0.061	0.145	0.076	0.095	0.079	0.067
% Other UGFT	-0.147	-0.008	0.032	-0.104	-0.068	0.059
UGFT Black	0.193	0.111	0.036	0.014	-0.067	0.058
Average loan of FTFT w FA	0.096	-0.009	-0.034	0.145	0.047	0.049
UGFT Alien	-0.189	0.005	-0.018	-0.025	-0.009	0.047
% FT Faculty Ranked	-0.060	-0.115	-0.086	0.065	0.079	0.023
% Hispanic UGPT	-0.271	0.007	0.030	0.115	0.058	0.012

Appendix A-6: The Student Achievement Model

The bachelor's degree-seeking model will report outcomes for (1) full time bachelor's degree-seeking students attending college for the first time (the same definition as the Student Right to Know graduation rate report in IPEDS); and (2) full-time bachelor's degree-seeking students who transferred in to the reporting institution (transfer-in students). Additionally, institutions have the option to report outcomes for (1) part-time bachelor's degree-seeking students attending college for the first time; and (2) part-time bachelor's degree-seeking students who transferred in to the reporting institution (transfer-in students).

The bachelor's degree-seeking model will report on the percentage of students who:

- Graduated from the reporting institution
- Are still enrolled at the reporting institution
- Transferred/graduated from one or more subsequent institution
- Transferred/are still enrolled at a subsequent institution
- Have unknown current enrollment or graduation status (includes students who transferred but whose enrollment or graduation status is unknown)

The bachelor's degree-seeking model will report outcomes for three time periods for each cohort.

- For full-time cohorts:
 - At the end of 4 years
 - At the end of 5 years
 - At the end of 6 years
- For part-time cohorts:
 - At the end of 6 years
 - At the end of 8 years
 - At the end of 10 years

<https://s3.amazonaws.com/studentachievementmeasure/SAM+Permanent+Links/SAM+Bachelors+Model+Methodology.pdf>