IMPROVING COLLEGIATE OUTCOMES AT BROAD-ACCESS INSTITUTIONS: LESSONS FOR RESEARCH AND PRACTICE

Michal Kurlaender
University of California, Davis

Jessica S. Howell
The College Board

Jacob Jackson
University of California, Davis

November 1, 2012

Improving Collegiate Outcomes at Broad-Access Institutions: Lessons for Research and Practice

Introduction

Today's broad-access colleges and universities are grappling with considerable pressures to improve collegiate outcomes, in particular, completion and time-to-degree. Although more people are entering college than ever before, degree completion has stagnated. College completion is particularly low in open-access or less selective postsecondary institutions, where time-to-degree has also been on the rise (Bound, Lovenheim, & Turner, 2012). Many of these institutions are struggling with what to do to improve completion rates among their students.

The purpose of this chapter is to outline how campuses can utilize data to identify and evaluate promising practices to improve college completion. We believe the principles and strategies we present are applicable to any number of goals postsecondary institutions may have, not just raising degree completion rates. As such, we hope the lessons are transferable and will serve to improve data-based decision making and leadership in broad access higher education institutions. Moreover, we do this using specific data examples from the nation's largest public higher education system—the California State University system. The CSU system, with 23 campuses, is the largest public higher education system in the country, educating about 1 in 10 California high school graduates and educating roughly 6 percent of the undergraduates enrolled in public four-year colleges in the entire nation.[1] Improving college completion has long

[1] This calculation is based on a published CSU enrollment of 450,000 students (http://www.calstate.edu/) and enrollment of 7.2 million student in public four year colleges nationwide in 2007 (http://nces.ed.gov/pubs2009/2009155.pdf).

been on the radar of CSU campuses.[2] CSU students come from urban, suburban and rural areas and attended public high schools that are both among the best and among the worst in the nation. While California may not be a typical state, it reflects well the student populations of other states in the U.S. and the mainstream public colleges that educate them.

The chapter is organized as follows. In section I, we provide some background and context to the college completion objective, emphasizing the important role of institutions in achieving these goals. In section II, we focus on quality data—how to get appropriate data, combine data from various sources and track students longitudinally in order to answer important questions about college success. In section III, we discuss how to analyze existing programs and policies, providing specific examples using institutional data from CSU. In section IV, we discuss how institutions might better evaluate newly implemented programs and efforts. Section V concludes by summarizing several key principles that are transferable to a host of other possible institutional goals.

I. Background/Context

Students who enter college fail to complete a degree for many reasons: loss of interest in college, lack of preparation or academic ability to persist, financial constraints, and/or institutional practices. First, not all students who enroll in college want to spend four years there. Students choose to enroll and subsequently complete a college degree based on an interaction of their preferences, academic ability, resources, and a variety of other factors. Thus, students regularly (i.e., each term when they register) have the

[2] The CSU system was awarded a competitive Complete College America state innovation challenge grant program: http://www.completecollege.org/path_forward/innovation_challenge/

opportunity to weigh the additional benefits of staying enrolled—such as increases in knowledge, potential earnings, and collegiate experiences—against additional costs like tuition, forgone earnings, and time spent in classes they dislike. If, at any point, this new information causes the perceived incremental costs associated with college to outweigh the incremental benefits of continuing, a student may choose to drop out or enroll part-time (at least temporarily).

Second, academic skills and preparation in high school are key predictors of college completion and success (Adelman, 2006; 1999; Dougherty, Mellor, & Jian, 2006; Fletcher & Tienda, 2008; Long, Conger, & Iatarola, 2012). It is no surprise that CSU freshmen with higher high school GPA and higher SAT scores are more likely to complete their college degree than their lower performing counterparts (Kurlaender, Jackson, & Howell, 2012). Many students arrive at college unprepared for college level work (Kurlaender & Howell, 2012; Snyder, Tan, and Hoffman, 2004) and those students who arrive at college in need of remediation are less likely to persist in college when compared to their peers who arrive better prepared academically (Bettinger & Long, 2009; Calcagno & Long, 2008; Boatman & Long, 2011; Martorell & McFarlin, 2011). At California State University, about 65 percent of students arrive in need of remediation in English and/or math as freshmen. CSU students identified in need of remediation in English and/or mathematics at entry are less likely to complete college than those who arrive at college prepared for college level work (Howell, Kurlaender, & Grodsky, 2010).

Third, financial constraints remain a barrier to college completion. Researchers have found direct evidence of the causal impacts of college costs and financial aid on college enrollment. Several studies have demonstrated that reducing college costs

increases college *enrollment* (Bound & Turner, 2002; Dynarski, 2003; Kane, 1994) especially among students with lower family incomes (Avery, Hoxby, Jackson, Burek, Poppe, & Raman, 2006). Income has become a more powerful determinant of college attendance over time, as well as of the quality of the colleges students attend (Belley & Lochner, 2007). Less is known about the causal impact of cost on college *completion*, but what we do know suggests that costs matter (Bettinger, 2004; Dynarski, 2005). Overall, this body of work suggests that financial constraints matter, however financial considerations beyond the direct costs of college require closer examination (Carneiro & Heckman, 2002; Stinebrickner & Stinebrickner, 2008).

Finally, institutional policies and practices may also play an important role in predicting degree receipt. Colleges vary widely with respect to the share of entering freshman they graduate within four, five or six years. What practices might account for institutional variation in rates of freshman completion and time to degree? Prior research suggests that student interaction with faculty, student peers and sense of community, active engagement with the institution, and mentoring all contribute to higher rates of persistence (Astin, 1993; Tinto, 1993; Lotkowski, Robbins, & Noeth, 2004). Although these provide promising directions for future research, many of these studies fail to adequately control for observable and unobservable differences between students who select different kinds of colleges or collegiate experiences (Astin, 1993; Braxton, 2000; Tinto, 1993) and thus likely conflate the contributions of student characteristics to institutional rates of postsecondary persistence with those of institutional practices. College selectivity accounts for an appreciable share of the institutional variation in college graduation overall (Small & Winship, 2007; Melguizo, 2008; Smith, 2012),

though work focusing specifically on community colleges has found less consistent evidence on the role of institutional quality measures on students' outcome (Stange, 2012; Calcagno et al., 2008; Sandy et al., 2006). More recently, several papers have suggested that cohort crowding and declining resources (particularly at less selective public institutions) may also lead to reductions in rates of college completion and increases in time to degree (Bound, Lovenheim, & Turner, 2010; 2012).

Although comparisons between colleges have shed light on the relative contributions of policies and student characteristics to graduation rates, most college campuses only have access to data on their own student body. It is likely that all colleges wish to improve degree completion rates and that individual campuses attract and admit fairly similar students from year to year. As such, colleges are uniquely suited to investigate the impacts of their campus-specific policies and programs, especially with more and more data being collected and reported at the local level.

In the example we use throughout this chapter, we analyze policies that may impact degree completion at a California State University campus. Specifically, we use student-level application, program participation, and degree data from one CSU campus to demonstrate methods of analyzing the impacts of three different programs on completion: a freshmen orientation program, declaring a major at college entry, and Summer Bridge.[3] We describe the analyses in detail within each section of this chapter.

[3] Freshman orientation programs seek to familiarize students with registration procedures, course offerings, academic requirements, and generally assist students with their adjustment to college. Analyses of important milestones and momentum points have identified the timely declaration of a major as an indicator of progress toward a degree (see, for example, Moore & Shulock (2011)). Summer Bridge programs are intensive, residential orientation programs for entering students who are not fully prepared to meet the demands of college coursework. Summer Bridge is one among many enrichment programs targeted towards first generation college students, or other students deemed "at risk" for college attrition or failure.

II. Quality Data

Data-driven decision making is now a ubiquitous component of K-12 school improvement. Unfortunately, this approach has not been fully transferred to higher education. College campuses often have better data on their alumni and potential donors than on the educational backgrounds of their existing students. There may be many reasons for data limitations in postsecondary schooling. In particular, higher education has not been the focus of intense accountability efforts as in K-12 and, as such, has had little incentive to invest in detailed data systems or the professional capacity necessary to maintain and analyze them. This may be changing with increased focus on college readiness as part of the Common Core State Standards movement and increased national policy attention on college attainment outcomes as a means of maintaining the competitiveness of the U.S. economy (Obama, 2009).[4] Moreover, higher education data systems are often highly decentralized, such that admissions offices collect information from student applications; financial aid offices collect only financial aid data; registrar offices collect only current registration information, and so on.

Collecting quality data is critical to addressing many of the most pressing issues facing higher education, including college completion. We present several important principles in assembling quality data to answer institutional questions. First, consider the source. This is essential for thinking about both data generalizability and data validity. Specifically, is the source providing the data (e.g., a survey given at summer orientation) representative of the general campus? And, is the data source a valid measure of what

[4] When President Obama declared that, "by 2020, America will once again have the highest proportion of college graduates in the world," a substantial number of college completion initiatives and efforts were already underway and more sprang up accordingly. For an overview of major U.S. college completion initiatives, see Russell (2011).

you are trying to capture? For example, students' self-reported data on family income is not as reliable of a source of information on students' potential financial constraints as is data available from the financial aid office that originates from the FAFSA.

Second, there are important advantages to combining data in order to provide a rich description of the student. Often, campus data sources can be separate, as each serves a function for one respective department or another. Application data may serve admissions, while degree data may be used by the registrar. Gathering all available data about a student means either streamlining collection and storage of data or merging the data together after the fact. In our example, we use data from four different sources within the same university. First, we use application data for information on detailed student characteristics. Second, we use term-by-term data to determine declared major at entry. Third, we use program participation data to determine orientation and summer bridge status. Lastly, we use degree data for a student's graduation status. It is only through combining these distinct data sources that we are able to more fully investigate the impacts of programs and policies on our student outcome of interest.

Third, build longitudinal data files to be able to more fully describe students' experiences prior to entering the institution. This is particularly important when those experiences may influence their postsecondary success. Many states (including California) do not yet have integrated data systems that allow researchers and education leaders to easily investigate students across the education pipeline. However, working separately with California's different education segments (i.e., K-12, community colleges, baccalaureate institutions), we have been able to assemble longitudinal data files to track California's high school students entering one of the of the State's 23

campuses of the CSU or one of the 112 California Community Colleges. In doing so, we have been able to investigate one of CSU's goal of improving college readiness among high school students (Kurlaender, Jackson, & Howell, 2012; Howell, Kurlaender, & Grodsky, 2009). For these studies we had to build a longitudinal data file connecting California high school juniors' school records with their later postsecondary outcomes at CSU and at the community college. It is also important that longitudinal data track students for as long as possible to determine if they ever obtain a degree at a different postsecondary institution and, importantly, how their postsecondary schooling experiences may have influenced later occupational outcomes. We do this in new work looking at the employment and earning outcomes of California's community college students (Kurlaender, 2012). Many states that already have an integrated data system have been able to do this more easily. Moreover, the importance of such an endeavor is underscored by the recent establishment of the Center for Analysis of Postsecondary Education and Employment, funded the U.S. Department of Education.[5]

III. Analyzing Existing Programs and Practices

Postsecondary institutions across the country are committed to improving student outcomes, in particular, degree receipt. Many have established new committees, programs and task forces addressing college attrition and delayed time-to-degree (see efforts such as Complete College America). CSU is no exception with a heavy system-wide focus on raising the six-year college completion rates, particularly for underrepresented minority groups. In this section, we offer some important guidance on

[5] See: http://capseecenter.org/

how institutions might effectively identify and analyze existing programs and practices that may improve college completion. Again, we note that the lessons here are transferable to a host of other institutional goals and objectives.

Principles for Identification

Campuses often begin their efforts by simply identifying promising programs, policies, and practices. We suggest that identification be based on a purposeful approach that may include one, if not many, of four simple principles. First, particular programs and/or policies should be based on evidence from the *existing literature* that is suggestive (if not convincing) that such programs, policies, and/or practices work, at least at other similar campuses.[6] Second, different programs, policies and practices come at varying costs to the institution, therefore institutional leaders should be aware of such *costs* in conjunction with their consideration of the benefits of existing programs. Third, institutions often *survey* their students for helpful insights about programs and policies that may aid, or possibly hinder, their academic progress. For example, student surveys might be helpful in streamlining registration processes, understanding obstacles to particular majors, or to provide a broad picture of student (dis)satisfaction with campus policies, procedures, and even programs. Finally, institutions can rely on their own administrative data and utilize an *empirically-based* approach to investigate existing programs' success and the impact of policies and practices. Ideally these approaches

[6] The quality of evidence in the research literature lies along a spectrum from correlational (suggestive) to causal (prescriptive). Because causal evidence is preferred but more challenging to generate, campus decision makers will likely have to rely on suggestive evidence in many instances. The What Works Clearinghouse (http://ies.ed.gov/ncee/wwc/) is an initiative of the U.S. Department of Education's Institute of Education Sciences (IES) that is designed to be a central and trusted source of scientific evidence for what works in education. This resource should always be examined for causal evidence on similar policies, programs, and practices under consideration on college campuses.

work together; that is, programs and policies often exist because they are based on what we know from the literature about student success, what we hear from student feedback via surveys, are transparent regarding differential cost, and rely on more regular quantitative evaluation of effects on measurable student outcomes.

Using the above criteria and rich data from one typical CSU campus, we now provide a concrete example of three such programs/practices that are often employed or considered at many postsecondary institutions. Specifically, we investigate this campus's freshmen orientation program, the practice of early declaration of a major, and participation in Summer Bridge. These three practices exist on most broad-access college campuses, and there is reason to believe that any or all of these programs could have an impact on college degree completion. They range from relatively low-cost (mandating that all students declare a major at entry) to relatively high-cost (expanding a Summer Bridge program).

Evaluating Existing Programs/Practices

Many campuses likely collect the necessary data to evaluate the effectiveness of existing programs, policies and practices. However, when such data does not exist, it is critical to have a data collection strategy that would yield reliable information about who participates in what programs and who might be implicated by particular policies.

We use data on two first-time freshmen cohorts attending this CSU campus in the 2002 and 2003 academic years and begin by examining who participates in these three programs. As Table 1 details, 60 percent of students at this CSU campus participated in

freshmen orientation, 75 percent of first-time freshmen had declared a major by the end of their first year, and 4 percent participated in the targeted Summer Bridge program.

Table 1: Pooled 2002 and 2003 First-time Freshmen Program/Practices

	Total N	Freshman Orientation	Declare Major	Summer Bridge
Number Participating	6,466	3,909	4,855	241
% Participating	-	60%	75%	4%

It may be helpful to begin by examining the outcomes of interest for program participants and non-participants. Such data can show immediately promising candidates for expansion or contraction, however such simple comparisons of outcomes can also be incredibly misleading and should not be relied upon without closer inspection, as we demonstrate below. Table 2 compares the six-year degree completion rates of these 2002 and 2003 first-time freshmen cohorts. Those who participate in freshmen orientation have a 56 percent completion rate compared to a 39 percent completion rate among non-participants (a difference of 17 percentage points). Among all freshmen who declare a major in their first year, 50 percent of students graduate within six years compared to 46 percent among those who declare their major later (a difference of 4 percentage points). Finally, among all first-time freshmen students who participate in Summer Bridge, 43 percent ultimately graduate compared to the 50 percent graduation rate among non-participants (a difference of -7 percentage points). Although the difference in graduation rates between Summer Bridge participants and non-participants is negative, suggesting the program may have the exact opposite effect of its intended goals, these raw differences may be misleading because of the important differences that distinguish program participants from non-participants across all of these programs (a critical point

we turn to next). Similarly, the 17 percentage point advantage between those who participate in freshmen orientation and those who do not may lead one to believe that the program is doing a great job at helping students to graduate, but orientation participants may be markedly different from non-participants in ways that skew this simplistic interpretation of the data.

Table 2: Pooled 2002 and 2003 First-time Freshmen Six-year Completion Rates by Program/Practices

	Freshman Orientation	Declare Major	Sumer Bridge
Number of students participating	3,909	4,855	241
% of Participants Completing Degree	56%	50%	43%
% of Non- Participants Completing Degree	39%	46%	50%
Difference	**17 ppts**	**4 ppts**	**-7 ppts**

Considering Selection into Program Participation

In general, students do not randomly decide to participate in programs or engage in particular practices. Some programs attract certain types of students and some are only targeted at specific student populations. Considering who is likely to participate in a program nearly always explains some of the relationship between program participation and student outcomes. This is crucial to consider when analyzing differences in student outcomes like those displayed in Table 2. In our example, recall that degree completion rates among freshmen orientation participants were 17 percentage points higher than non-participants. In the top panel of Table 3, we present information about the characteristics of those who participate in freshmen orientation and those that do not, noting where those differences are statistically significant. Students who participate in freshmen orientation are quite different from those who do not; specifically, they are more likely to be white or

Asian, female, exempt from needing remediation in Math and English, respectively, to have a higher high school grade point average and SAT score, and are less likely to be from a lower income household or to have parents with a high school diploma or less.

Although only 25 percent of first-time freshmen do not declare a major by the end of their first year, the second panel of Table 3 indicates that those students who wait to declare a major appear to have significantly weaker academic credentials relative to those who do declare a major within their first year (e.g., more likely to need remediation in English and/or Math, lower high school GPA and lower SAT scores), and significantly more likely to be first-generation college students. Finally, we note that Summer Bridge is (as intended) a highly targeted program and, as such, students who participate in Summer Bridge are significantly different from those who do not along a host of dimensions. In particular, the third panel of Table 3 indicates that Summer Bridge students are more likely to be from underrepresented minority groups (Latino or African American), to be female, to have lower academic credentials (save English remediation), and more likely to be from a lower socio-economic background (based on family income and parental education levels).

Table 3: Mean Differences in Student Characteristics by program

Panel 1: Freshman Orientation

	Yes	No	Diff	
URM	0.31	0.42	-0.11	***
Male	0.38	0.41	-0.03	**
Math Exempt	0.78	0.63	0.15	***
English Exempt	0.57	0.44	0.13	***
HS GPA	3.28	3.14	0.15	***
SAT	994	948	46.68	***
Low Income	0.19	0.31	-0.12	***
Missing Income	0.10	0.10	0.00	
First Generation	0.17	0.28	-0.11	**
Missing Parent Ed	0.07	0.08	-0.01	**
Disability	0.02	0.02	0.00	

Panel 2: Declare Major

	Yes	No	Diff	
URM	0.36	0.35	0.01	
Male	0.39	0.40	0.00	
Math Exempt	0.73	0.68	0.05	***
English Exempt	0.53	0.47	0.06	***
HS GPA	3.24	3.18	0.06	***
SAT	981	962	18.71	***
Low Income	0.24	0.23	0.01	
Missing Income	0.10	0.10	0.00	
First Generation	0.21	0.23	-0.02	**
Missing Parent Ed	0.07	0.08	-0.01	
Disability	0.02	0.01	0.01	

Panel 3: Summer Bridge

	Yes	No	Diff	
URM	0.74	0.34	0.40	***
Male	0.32	0.40	-0.07	**
Math Exempt	0.62	0.72	-0.10	***
English Exempt	0.71	0.51	0.20	***
HS GPA	2.89	3.24	-0.35	***
SAT	839	981	-142.24	***
Low Income	0.63	0.22	0.41	***
Missing Income	0.10	0.10	0.00	
First Generation	0.58	0.20	0.37	***
Missing Parent Ed	0.07	0.08	0.00	
Disability	0.02	0.02	0.00	

It is critical to account for these observable student and family differences when evaluating the impact of these programs/practices on student outcomes for two primary reasons. First, the student characteristics associated with participating in freshmen orientation or Summer Bridge and for having declaring a major early are the same characteristics that are associated with college completion. In other words, first generation college students may not only be less likely to attend freshmen orientation than students whose parents have higher educational attainment, but first generation students are also less likely to finish college more generally. Similarly, students who need remediation in college may be less likely to declare a major in their freshmen year than those who do not need remediation, and so it is hard to disentangle whether it is the fact that they didn't declare a major in their first year or their lack of academic readiness that keeps them from graduating at the same rate as those who do declare a major early in their college career. Of course, there are likely also a host of other unobserved characteristics that may be associated with college completion, such as student motivation, that may also be associated with selection into these programs. It is much more difficult to account for factors that we do not observe, but it is nonetheless important to consider what these factors might be and how they probably interact with both program participation and our outcome—college completion.[7]

If participation in a program is conflated with the effectiveness of the program, we need to attempt to separate those effects in our analysis. This is accomplished, to a

[7] For example, relatively more motivated students are likely the ones who participate in freshman orientation when it is optional and their motivation also arguably propels them to greater success in their coursework and in completing a degree. Although it is very challenging to quantify the impact of motivation, we note that the direction of its effect is clear, allowing a clear understanding of bias associated with mandating freshman orientation. If those students who are induced to participate in freshman orientation when it becomes mandatory are relatively less motivated students, then the impact of the mandate on student academic outcomes will be mitigated by this unobservable factor that we know exists but cannot directly incorporate into the analyses.

large extent, by accounting for the characteristics of participants in a regression analysis framework.[8] These "regression adjusted" differences in completion rates for participants compared to non-participants are presented in Table 4.

In our example, the raw differences in completion rates between students who participated in freshmen orientation and those who did not (17 percentage points in Table 2) is substantially reduced to an 11.5 percentage point difference when we account for student characteristics associated with orientation participation in Table 4. Similarly, when we control for differences in student characteristics in our comparison of freshmen students who declare a major by the end of their first year with those who do not, the graduation rate among those who declare a major early is only 2.4 percentage points higher than those who do not (a difference that is not only smaller than the 4 percentage point raw difference in Table 2, but is now no longer statistically different from zero, Table 4). Finally, when we adjust for the characteristics of students who are selected into the Summer Bridge program, we now find a positive relationship with college completion. Specifically, among similar students who meet the Summer Bridge criteria, participating students have a significantly higher college graduation rate than those who do not participate in Summer Bridge (a difference of 8.7 percentage points in Table 4 compared to the misleading negative estimated effect in Table 2).

[8] Regression analysis is a statistical technique for quantifying the relationships between variables that are observable in data. In multiple regression, the effect of changing one variable (e.g., high school GPA) on the outcome of interest (e.g., college completion) is isolated by holding constant other variables that also influence the outcome (e.g., race/ethnicity, parental education and income, etc.). Regression analysis, although usually conducted by researchers using specialized statistical software, is also available in Excel using the Data Analysis Toolpack add-in.

Table 4: Regression Adjusted Program/Practices Effects

	Freshman Orientation	Declare Major	Summer Bridge
Number of students participating	3,909	4,855	241
% of Participants Graduating	56%	50%	43%
% of non-Participants Graduating	39%	46%	50%
RAW Difference	17 ppts	4 ppts	-7 ppts
Regression Adjusted Difference	**11.5 ppts*** **	**2.4 ppts**	**8.7 ppts***

The second reason it is so critical to think about who participates in particular programs is because it provides campuses with vital information about how their specific student population may respond to changes in programs, policies and practices. In other words, campuses need better information about the characteristics of their students who engage in particular activities in order to assess whether altering those programs is likely to yield desired changes or have unintended consequences. Having such detailed information about students also allows for useful forecasts into the future. Campuses can project what might happen to completion rates given changing demographics of the student population.

Simulating Changes in Policy

Adjusting program effects to account for student characteristics and selection into program participation can help identify which programs may have favorable impacts and for what types of students. In addition, campuses can use similar methods of analysis (based on student characteristics and program participation) to project student outcomes in the future and to simulate what might happen to student outcomes if a particular campus program were expanded or cut. The "2025 California Goal Dashboard" is a good example of a prediction tool that is a more advanced version of the reporting tools that

are increasingly known as "data dashboards."[9] Such prediction tools harness regression modeling techniques and student-level data to forecast what might happen to various outcomes of interest when we change some of the inputs.

In our example in this chapter, we have used completion data from 2002 and 2003 to build a model that suggested that students who participated in a freshmen orientation program on a specific CSU campus were 11.5 percentage points more likely to graduate than students who did not, even after adjusting for important demographic and academic characteristics of students who participate. We can use those same models to simulate what would have happened if freshmen orientation had been made mandatory on that campus, causing all students to participate in the program. First, we apply the model generated by our regression analysis in Table 4 to the pooled 2002 and 2003 data with true information on program participation to generate a baseline projection. As expected, the model quite accurately predicts the completion rate of the pooled 2002 and 2003 cohorts at 49.4 percent in Table 5 (with the actual graduation rate at 49.2).

Next, we simulate what would have happened to graduation rates if everyone in the 2002 and 2003 freshmen cohorts participated in freshmen orientation. To do so, we generate a modified dataset in which each student appears to have participated in freshmen orientation. We apply the same model parameters from our regression analysis to the modified data and calculate the expected increase in completion rates on this campus (also presented in Table 5). The results from our simulation suggest that we

[9] Other states (e.g., Virginia; see http://research.schev.edu/default.asp?select1=Reports) and the U.S. Department of Education (see http://dashboard.ed.gov/dashboard.aspx) also have data dashboards that report the sort of data that could be harnessed for the type of simulation and prediction available on the 2025 California Goal Dashboard (which is available at http://californiacompetes.weba.autoupdate.com/flash/ca_studentflowmodel.swf).

would expect a graduation rate near 54 percent for these cohorts, which is a predicted increase in the college completion rate of roughly four to five percentage points.

Table 5: Baseline and Simulated Graduation Rates for Pooled 2002 & 2003 Cohorts

Baseline Graduation Rate (predicted based on existing cohort characteristics)	49.4%
Simulated Graduation Rate (based on 100% freshman orientation participation)	53.9%
Projected Gain in Completion from Policy Change	**4.5 ppts**

Why does the simulation project only a 4.5 percentage point gain in the completion rate rather than the 11.5 percentage point gain the campus may have hoped for given the regression adjusted program effects presented in Table 4? The simulated completion rate is substantially lower simply because most students already attend freshmen orientation, and those that do not have a variety of characteristics associated with lower odds of graduation, including lower SAT and high school GPA (see Table 3). It is also likely, though nearly impossible to demonstrate with data, that the non-participants induced into orientation by such a policy have unobservable characteristics that are also associated with lower odds of graduation (e.g., lower motivation). Mandating the participation of those additional students cannot offset the negative influence of their other characteristics on completion.

We could apply a similar simulation technique to early major declaration and/or Summer Bridge participation. A great majority of students come into the institution with a declared major, but campuses could use the methodology described above to predict what would happen if the campus made this a mandatory part of enrollment; as suggested above we would need to consider what might be different between those who do and do

not choose to declare a major at entry. And, although it may not be sensible to simulate universal participation in Summer Bridge (since it is such a small and targeted program), we can, however, simulate what would happen if simply more students with the specific characteristics targeted by Summer Bridge (first generation college students, often from under-represented minority groups) participated in Summer Bridge. By narrowing the simulation to the population likely implicated in expanding a narrow program such as Summer Bridge, campuses can get a more realistic picture of the potential effects of such action.

Expanding Programs: A Cautionary Tale?

In 2004, the CSU campus we have been using for our example decided to make freshmen orientation a mandatory program. This policy change provides us with a unique opportunity to also examine the impact on college completion rates of an actual (rather than simulated) policy change and to compare our simulation model to actual data. As noted previously, about sixty percent of freshmen attended orientation in the combined 2002 and 2003 freshmen cohorts. In 2004, when the campus made orientation mandatory, nearly all students (96 percent) participated. Other than orientation participation, there are few changes in student characteristics across the 2002, 2003, and 2004 cohorts. If our within-sample simulations in Table 5 showed a 4.5 percentage point increase in graduation rates by making orientation a mandatory program, we would expect to see a similar increase in graduation rates for the 2004 cohort, when the program was actually formally expanded.

What happened to the graduation rates of the 2004 freshmen for whom orientation was mandatory? The results are shown in Table 6. For the 2004 cohort, our simulation method predicts a graduation rate of 54.2 percent, which is a nearly 5 percentage point projected increase above either simulated or actual 2002/2003 completion rates. Because this policy change was actually implemented, however, we know that the six-year graduation rate for the 2004 cohort was essentially unchanged at 49.5 percent despite the mandatory orientation policy.

Table 6: Projected Versus Actual Six-Year Graduation Rates Following Mandatory Freshman Orientation Policy in 2004

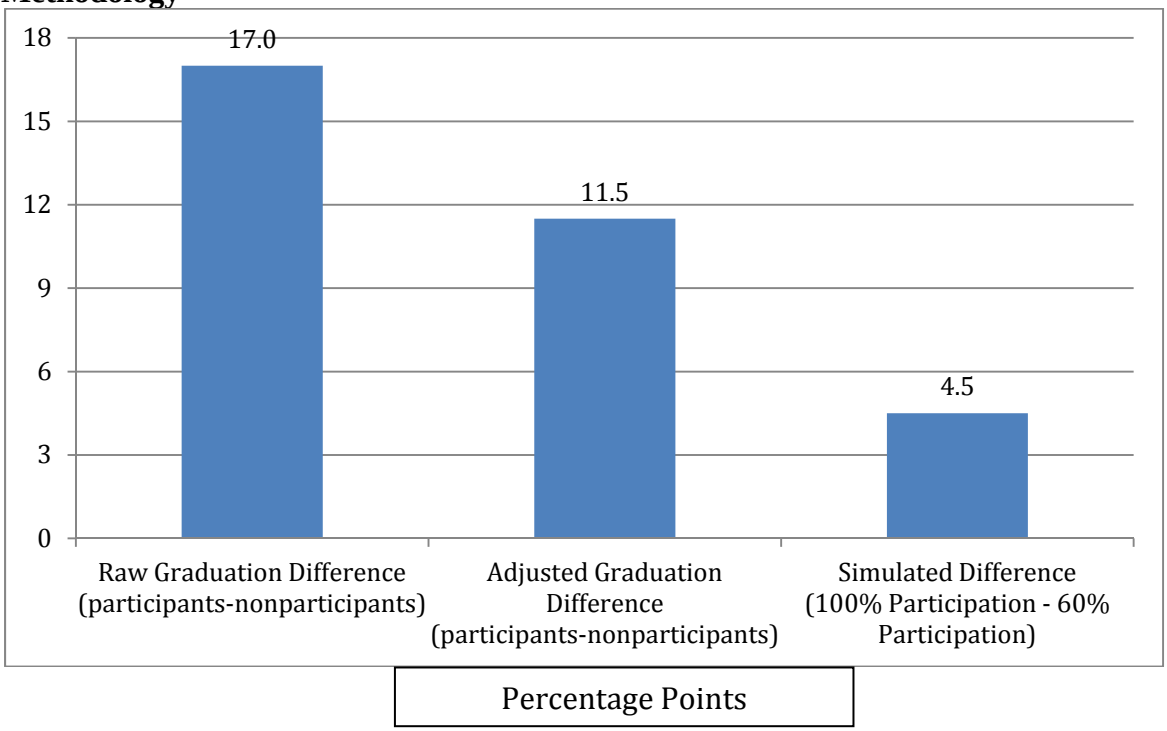
Cohort	2002/2003	2004
Orientation Participation	60.7%	96.3%
Simulated Graduation Rate	49.4%	54.2%
Actual Graduation Rate	49.2%	49.5%

Why didn't universal participation in freshmen orientation lead to the expected improvement in graduation rates? Our simulation model tells us it is not because of demographic or academic changes in the students, since we account for these differences in our model. So, what went wrong? First, is nearly impossible to isolate this program change from other changes experienced by the university (e.g., changes in administration, economic conditions or other resources, etc.). Moreover, although our model accounts for the many observable characteristics associated with program participation and graduation, there are likely a host of other unobserved characteristics that may be associated with college completion, such as motivation, that may also impact graduation rates. Finally, it is possible that, in the expansion of the orientation program to all students, the program was less effective because expansion changed the program in some

way (e.g., led to weaker instructors for freshmen orientation or a watered down freshmen orientation experience). In all likelihood, a combination of several of these factors is likely at play, resulting in impacts that are smaller than expected and drawing attention to the importance of appropriately implementing the scale-up of programs and practices.

Although the simulation model did not match the actual data on completion rates following this policy change, it is important to reiterate the value of this methodology for making more accurate predictions than would have otherwise been available to campus decision makers. Figure 1 summarizes this important take-away. If campus decision makers simply compared raw completion rate differences between participants and non-participants, they would have naively expected a mandatory orientation policy to boost completion rates by 17 percentage points. If these same decision makers were more data savvy and thought to adjust the raw differences using regression analysis and the student characteristics associated with participation, they would have expected a mandatory orientation policy to boost completion rates by 11.5 percentage points. Finally, if decision makers used their regression results to simulate the effect of a mandatory orientation policy, they would have expected a 4.5 percentage point increase in degree completion. Although all three methods overstate the true impact, observed to be zero in this case, the simulation method demonstrates the greatest potential to remove factors that are known to bias upwards the anticipated benefits of the policy change. We argue that improved accuracy is a worthy goal when perfect accuracy is not possible.

Figure 1: Differences in Anticipated Benefits of Mandatory Orientation, by Methodology



IV. New Directions for Institutional Program Evaluation

In this final section, we discuss how postsecondary institutions can employ some basic and critically important principles in implementing and evaluating programs and policies. Specifically, we focus on ensuring that programs implemented across college campuses can be grounded in existing literature about what works, a thoughtful cost-benefit analysis, and a more rigorous evaluation plan to identify program effects with improved accuracy.

Campuses do not necessarily have to reinvent the wheel when deciding what new programs and policy changes to initiate. As suggested above, programs should be grounded in both existing literature about what works, and with the current realities of the institution in mind. We suggest that a close investigation of institutional data is also important for identifying promising practices and programs. First, how similar are the

student characteristics of your campus to those discussed in other settings where programs and policies were deemed successful? Second, institutions can use existing data to explore how many students would be implicated by a new program, which may be targeted (such as Summer Bridge), or a policy change, such as establishing a new minimum credit load. All of this information is essential for both the design of a successful program and for subsequently understanding why programs may or may not have met their intended goals.

Another critical step in the process of identifying promising programs for expansion or removal is a cost-benefit analysis.[10] This chapter has identified a method of projecting the benefits of policy and programmatic changes, but those benefits must be put on a "per dollar" basis using project costs in order to appropriately compare alternatives. Although an analysis of the direct costs of expanding a program is beyond the scope of this chapter, individual campuses can probably obtain the necessary information on the direct costs of programs more readily (and perhaps from campuses that have already implemented such changes). The real strength of the cost-benefit analysis is that campus decision makers are then able to weigh the impact on their degree completion rates of an additional dollar spent various different ways. Imagine the power of knowing how an additional dollar invested in freshman orientation, an initiative to declare majors in the first year, or Summer Bridge is differentially predicted to influence completion rates. On larger campuses, larger sample sizes would even enable decision makers to examine how to move the completion needle for particular sub-groups of

[10] A good source on cost-benefit Analysis in education is: "*Cost-Effective Analysis: Methods and Applications,*" by Henry Levin and Patrick McEwan (Sage Publications, Inc.).

interest, like first generation students, English language learners, ethnic minorities, and so on.

Institutions, often with good intentions to solve existing problems, frequently undertake major changes or implement new programs based on limited information. We offer a strong suggestion to pilot new programs and/or policy changes when at all possible. This provides valuable information about how new programs (or changing policies) may function on the ground, who is impacted, and the intended and unintended consequences of such changes. Although the focus of this chapter has been on quantitative evidence, campuses should also collect qualitative information during the pilot phase on the specifics of implementation, which will be crucial if the results of the pilot lead the campus to scale up to a larger portion of the student body.

Once campuses make a decision to implement or expand a new program, or to alter an existing policy or practice, it is vital to establish an appropriate comparison group in order to effectively evaluate the change.[11] The examples we provide in the previous section yield some initial insights about why this is important. Mainly, program participants are significantly different from non-participants in both observable and unobservable ways, either because some programs are targeted toward certain students (like Summer Bridge for first-generation, underrepresented minority students) or because programs are optional (freshmen orientation in 2002 and 2003). As a result, participation in such programs is not random and this lack of randomness implies that there may not be

[11] See "*By Design: Planning Research on Higher Education*,"by Richard Light, Judith Singer, and John Willett for an excellent discussion of this.

a simple control group available for comparison without some proactive design prior to implementing changes. [12]

Randomized control trials are the only way to establish a true causal relationship between programs and outcomes. It is only by randomly assigning some students to be program participants and others to not be that we can confidently conclude whether observed changes in student outcomes are the result of participation in the program. Randomized experiments are difficult to accomplish in higher education and are often not feasible for a variety of reasons. Campus leaders are often adamantly against depriving students of a particular experience if there is enough room to accommodate all those interested and sufficient anecdotal evidence that the program will improve student outcomes. Even when there are program constraints, a first-come first-serve basis often feels like the most reasonable way to distribute resources; yet, those students who volunteer to participate first are potentially substantially different from those who do not, and we have demonstrated in this chapter that self-selection by students biases the interpretation of the data in ways that challenge the goal of rigorous data-driven decision making. Finally, changes in procedures or policies are nearly impossible to assign to some students and not to others in a campus environment, making the separation of treatment and control groups a very difficult hurdle.

In the absence of randomized experiments, it is useful to engage in some basic counterfactual thinking in order to design a solid evaluation plan (i.e. what would college completion look like in the absence of this program or change in policy?). Among social science researchers, these sorts of evaluations are described as quasi-experimental. Over

[12] We suggest the following user-friendly publication, which describes many of these issues in a non-technical way: Identifying and Implementing Educational Practices Supported By Rigorous Evidence: A User Friendly Guide (Available at: http://ies.ed.gov/ncee/pdf/evidence_based.pdf.)

the past two decades there have been important developments in education research and related fields in the use of "quasi-experimental" methods for program evaluation; these techniques aim to approximate the underlying logic of the experiment without random assignment (Murnane and Willett, 2011). These methods rely on finding comparisons that may exist naturally rather than being assigned to them randomly (Shavelsom & Towne, 2002). For example, a natural comparison group to evaluate the impact of a universal change in the minimum credit load on time-to-degree may be the cohorts of students enrolled right before the change in policy was implemented (assuming there are no other changes that occurred simultaneously that may have also influenced time-to-degree). Similarly, in evaluating the effects of collegiate remediation, many scholars compare students just above and just below the score cutoff for remediation placement, hypothesizing that at some level the cutoff is arbitrary (random) for students in the narrow band right around it. These approaches are fundamentally about eliminating the potential influence of observable and unobservable factors that may be confounded with the program effect (or the policy/program change).

V. **Conclusion and Implications for Practices**

Today, more than ever, colleges and universities are focused on improving persistence and degree completion for their students. Thus, it is imperative that such efforts be guided by some basic principles of design and evaluation if postsecondary institutions are to reap the benefits of the investment. First, postsecondary institutions need to have better knowledge of their student body. Campuses collect a large amount of information from students in admission about students' prior academic performance,

financial aid forms about student socioeconomic status, and enrollment practices (e.g., declaration of major, on-time degree accumulation, major, and academic performance). This information is useful for providing richer and more detailed data about the student body, and of particular subgroups of interest (e.g., first generation students, part-time students, etc.) by the desired outcomes—in this case, persistence, degree completion, and time-to-degree. In obtaining greater knowledge about the student body, campuses can more adequately identify differences among sub-populations to target and which practices may be more effective.

Second, campuses seek to identify promising practices and interventions. In so doing, they can *initiate* new programs or policies (e.g., offer a freshmen experience course), *expand* an existing one (e.g., make a voluntary program such as freshmen orientation mandatory), or *reduce or eliminate* a practice or policy that may present obstacles for students (e.g., reducing the number of required general education credits). In each of these cases it is critical for campuses to consider which students are affected by such a policy or practice change, and by what mechanism should the proposed change lead to improved outcomes.

Finally, campuses must be diligent in monitoring and evaluating the effects of new practices in order to determine if they have had the desired effects on student outcomes. In so doing there are some basic questions that are useful to consider:

Who is affected?

- Who currently participates in such programs or activities and are they the "typical" student?
- If creating a new program, will it be voluntary and then who might likely participate?
- What are the differences between participants and non-participants in background characteristics and in their enrollment practices?

- Is there any enforcement regarding who participates? Are there any consequences for not participating?

Change by what Mechanism?

- Why might the change in policies or programs lead to desired effects?
- What are the specific inputs of resources (e.g., smaller classes, mentoring program, priority registration, etc.) that are to alter outcomes?
- Is there any evidence of compliance for the program and/or policy change?

How to track change?

- What did these student outcomes look like before the new policy or program came into place?
- Did any other policies or programs get initiated or changed at the same time, and which of these may be conflating the changes observed in the data?
- What else changed about the institution and the student body that may either magnify or reduce the desired changes brought about by a new program or policy?

In closing, we encourage campus leaders and decision makers to recognize that they have potentially untapped resources on their campuses that may be of great help in implementing the types of analyses described in this chapter. It is well known that institutional research offices, particularly at broad-access institutions, are understaffed and existing resources may be exhausted by federal and state reporting requirements. All campuses, however, employ a wide variety of quantitatively-inclined instructional faculty in their departments of economics, education, public policy, sociology, and statistics. Social science researchers, particularly those looking to engage their particular skills in ways that serve the campus community, are ripe for engagement in data-driven decision making like what we outline here. We encourage campuses to think outside of the box and remove the silos that exist between the academic and student services sides of their

campuses in order to embrace the methods outlined in this chapter with the resources they already have at their disposal.

References

Adelman, C. (1999). *Answers in the tool box: Academic intensity, attendance patterns,and bachelor's degree attainment*. Washington, DC: U.S. Department of Education.

Adelman, C. (2006). *The Toolbox Revisited: Paths to Degree Completion From High School Through College*. Washington, DC: U.S. Department of Education.

Astin, A. (1993). *What matters in college? Four critical years revisited*. San Francisco: Jossey Bass.

Avery, C., Hoxby, C., Jackson, C., Burek, K., Pope, G., & Raman, M. (2006). *Cost should be no barrier: An evaluation of the first year of Harvard's financial aid initiative* (No. w12029). National Bureau of Economic Research.

Belley, P., & Lochner, L. (2007). *The changing role of family income and ability in determining educational achievement* (No. w13527). National Bureau of Economic Research.

Bettinger, E. (2004). How financial aid affects persistence. In *College choices: The economics of where to go, when to go, and how to pay for it* (pp. 207-238). University of Chicago Press.

Bettinger, E., & Long, B. T. (2009). Addressing the needs of underprepared students in higher education: Does college remediation work? *Journal of Human Resources, 44*(3), 736–771.

Boatman, A., & Long, B. T. (2010). *Does remediation work for all students? How the effects of postsecondary remedial and developmental courses vary by level of academic preparation* (Working Paper). National Center for Postsecondary Research.

Bound, J., Lovenheim, M. & Turner, S. (2012). Increasing Time to Baccalaureate Degree in the United States, *Education Finance and Policy*, 7(4), 375-424 (doi:10.1162/EDFP_a_00074)

Bound, J., Lovenheim, M. & Turner, S. (2010). Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources, *American Economic Journal: Applied Economics*, American Economic Association, 2(3), 129-57.

Braxton, J. M. (Ed.). (2000). *Reworking the student departure puzzle*. Vanderbilt University Press.

Calcagno, J. C., Bailey, T., Jenkins, D., Kienzl, G., & Leinbach, T. (2008). Community college student success: What institutional characteristics make a difference?. *Economics of Education Review*, *27*(6), 632-645.

Calcagno, J. C., & Long, B. T. (2008). *The impact of postsecondary remediation using a regression discontinuity approach: Addressing endogenous sorting and noncompliance* (NBER Working Paper No. 14194). Cambridge, MA: National Bureau of Economic Research.

Carneiro, P., & Heckman, J. J. (2002). The Evidence on Credit Constraints in Post-Secondary Schooling. *The Economic Journal*, *112*(482), 705-734.

Dougherty, C., Mellor, L., & Jian, S. (2006). The relationship between Advanced Placement and college graduation. *National Center for Educational Accountability*.

Dynarski, S. (2005). *Building the stock of college-educated labor* (No. w11604). National Bureau of Economic Research.

Dynarski, S. (2003). "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion." *American Economic Review*, 93(1): 279–288.

Fletcher, J. M., & Tienda, M. (2009). High school classmates and college success. *Sociology of Education*, *82*(4), 287-314.

Howell, J., Kurlaender, M., & Grodsky, E. (2010). Postsecondary preparation and remediation: Examining the effect of the Early Assessment Program at California State University. *Journal of Policy Analysis and Management*, *29*(4), 726–748.

Kane, T. J. (1994). College entry by blacks since 1970: The role of college costs, family background, and the returns to education. *Journal of Political Economy*, 878-911.

Kurlaender, M. & Howell, J. (2012). College Remediation: A Review of the Causes and Consequences. *Advocacy & Policy Center Literature Brief*. New York, NY: The College Board.

Kurlaender, M., Jackson, J., & Howell, J. (2012). K-12 Postsecondary Alignment and School Accountability: High School Responses to California's Early Assessment Program. *Advocacy & Policy Center Research Brief*. New York, NY: The College Board.

Light, R. J., Singer, J. D., Willett, J. B. (1990). By Design: Planning Research on Higher Education. Harvard University Press, Cambridge, MA.

Long, M., Conger, D. & Iatarola, P. (2012). "Effects of High School Course-Taking on Secondary and Postsecondary Success." *American Educational Research Journal*, 49(2), 285-322.

Lotkowski, V. A., Robbins, S. B., & Noeth, R. J. (2004). The role of academic and non-academic factors in improving college retention. *ACT Policy Report*, 20-24.

Martorell, P., & McFarlin, I. (2011). Help or hindrance? The effects of college remediation on academic and labor market outcomes. *Review of Economics and Statistics*, *93*(2), 436–454.

Melguizo, T. (2008). Quality matters: Assessing the impact of attending more selective institutions on college completion rates of minorities. *Research in Higher Education, 49*(3), 214-236.

Moore, C. & Shulock, N. (2011). Sense of Direction: The Importance of Helping Community College Students Select and Enter a Program of Study. *Institute for Higher Education Leadership & Policy*, Sacramento, CA. Retrieved from http://www.csus.edu/ihelp/PDFs/R_Sense_of_Direction.pdf.

Obama, B. (2009, February). Remarks of President Barack Obama — Address to Joint Session of Congress, White House Press Office, Washington, DC.

Russell, A. (2011, October). A Guide to Major U.S. College Completion Initiatives. *Higher Education Policy Brief, American Association of State Colleges and Universities.* Retrieved from http://www.aascu.org/policy/publications/policymatters/2011/collegecompletion.pdf.

Sandy, J., Gonzalez, A., & Hilmer, M. J. (2006). Alternative paths to college completion: Effect of attending a 2-year school on the probability of completing a 4-year degree. *Economics of Education Review, 25*(5), 463-471.

Shavelson, R. J. and Towne, L., Eds. (2002). Scientific Research in Education: Committee on Scientific Principles for Education Research, National Research Council. National Academy Press, Washington, DC.

Small, M. L., & Winship, C. (2007). Black Students' Graduation from Elite Colleges: Institutional Characteristics and Between-Institution Differences. *Social Science Research, 36*(3), 1257-1275.

Smith, J. (2012). Ova and Out: Using Twins to Estimate the Educational Returns to Attending a Selective College. Paper presented at the 2012 APPAM Fall Research Conference.

Stange, K. (2012). Ability Sorting and the Importance of College Quality to Student Achievement: Evidence from Community Colleges. *Education Finance and Policy*, 7(1).

Stinebrickner , T.& Stinebrickner, R., (2009). Learning about Academic Ability and the College Drop-out Decision, NBER Working Papers 14810, National Bureau of Economic Research, Inc.

Stinebrickner, T. & Stinebrickner, R. (2008). Learning About Academic Ability and the College Drop-Out Decision, University of Western Ontario, CIBC Centre for Human Capital and Productivity Working Papers 20086, University of Western Ontario, CIBC Centre for Human Capital and Productivity.

Tinto, V. (1993). Leaving College: Rethinking the causes and cures of student attrition, 2nd ed. Chicago: University of Chicago Press.

Turner, S. & Bound, J. (2003). Closing the Gap or Widening the Divide: The Effects of the G.I. Bill and World War II on the Educational Outcomes of Black Americans, The Journal of Economic History, Cambridge University Press, vol. 63(01), pages 145-177, March.