Data Collection and Use at Community Colleges

Peter T. Ewell
National Center for Higher Education Management Systems (NCHEMS)

Background

Community colleges are crucial to attaining President Obama's goal of placing the United States once again at the forefront of the world with respect to collegiate attainment. They enroll more than half the undergraduate students in the country. And not only do they produce workforce-related credentials and degrees in their own right but, by preparing hundreds of thousands of students to transfer, they play an important role in baccalaureate degree production as well.

But community colleges are also extraordinarily complicated environments with respect to the way students move into and through them. Students frequently interrupt their enrollment by stopping in and stopping out. They change programs a lot, and make use of myriad support services such as tutoring and supplemental instruction. More than half of them begin with developmental studies in reading, writing, or mathematics, because they are assessed below the college level in these skills. Large numbers also delay taking important "gatekeeper" courses or take courses out of sequence. Community college academic leaders and faculty need to deeply understand these factors in order to design and improve curricula and academic interventions, and they need appropriate and timely early warning systems to detect students in trouble. These kinds of enrollment "swirls" are equally prominent among community colleges and four-year institutions. According to federal studies, more than two-thirds of students who ultimately earn a baccalaureate degree attended two or more institutions to do so and one in five attended three or more (Adelman 2006). So a parallel understanding of student flow is mandatory for policymakers at the state or system level in order to improve attainment.

Longitudinal Data Systems

At both the institutional and state/system levels, the data needed to answer these kinds of questions can only be produced by longitudinal databases constructed on a cohort basis. These databases are built up from established student registration records to track groups of students who enter the institution (or state, or system) at the same time over multiple years of enrollment to determine their educational experiences and ultimate degree attainment. To be effective as information resources, moreover, longitudinal databases need to be extremely flexible—able to break down student attainment or performance measures by many combinations of student characteristics—male Hispanic students on Pell assistance seeking an automotive technology certificate or female students aged 25-34 entering with below-college mathematics skills seeing an associate degree in Nursing, for example. They also must contain a significant amount of "treatment" data that encompasses various aspects of the student experience that are presumed to be related to academic success—performance in "gate-keeper" courses like English Composition or a first college-level mathematics course, developmental placement and academic "catch-up" experiences, participation in tutoring and academic support, receipt of financial assistance in various packages, etc. Crossing these two kinds of data, academic administrators can begin to understand what works for whom, a fundamental condition for systematic improvement. For example, this kind of fine-grained tracking at Valencia Community College helped the college design improvements in the developmental mathematics sequence that increased success rates for African American and Hispanic students by between eight and ten percent in just a few years (Finney and Stoel, 2010).

Frequently overlooked by institutional actors in this array of resources are multi-institutional databases capable of tracking students after they leave a given community college. According to the latest inventory, "student unit record" (SUR) databases of this kind now exist in 45 states (Garcia and L'Orange, 2010). These can document successful transfer, and some can further examine student

performance at transfer institutions. To supplement state SUR databases, the National Student Clearinghouse (NSC) maintains enrollment and completion data on well over 90% of students enrolled in postsecondary education in the country. Finally, some 23 states have linked their SUR data to employment databases held by their workforce agencies to determine job placement and earnings. These external data resources are generally beyond the reach of individual institutions, but the fact that most community colleges are part of a system means that the data are accessible through system offices.

Performance and Outcome Measures

While completing a credential or transferring successfully to a college or university at a higher level are the ultimate success measures for community colleges, there is growing consensus that intermediate outcome measures are important in marking the progress of different kinds of students. For example, the Bill and Melinda Gates Foundation is currently developing a list of such measures that includes successful completion of developmental study for students assessed as not college ready, passage of gatekeeper courses in English and Mathematics, persistence to the next year, the attainment of several college-credit accumulation milestones (12, 24, and 42 credits), and the ratio of courses enrolled for that are completed with a grade of "C" or better—in addition to earning a college credential, transferring to a four year institution, earning a degree from the transfer institution, and employment in field. Similar measures are already in use by a consortium of states and are the basis of the exemplary "Momentum Points" performance funding scheme now operated by the State Board of Community and Technical Colleges (SBCTC) in Washington. Such intermediate measures enable the progress of an entering student cohort to be tracked much more finely—a substantial virtue for community college populations that frequently get into academic trouble early in their college careers.

Non-Credit Programs and Populations

While most longitudinal database development has taken place within the relatively traditional universe of academic degree credits (including developmental), both policymakers and college leaders are recognizing the growing importance of bringing the large non-credit student populations—ABE, GED, and occupational/vocational certificate enrollments—into the tracking universe. They realize that in order to attain the nation's ambitious attainment goals, success rates will need to improve for all students who are seeking to advance, regardless of their current level. For example, the common Gates Foundation measures define entering cohorts to include such students and propose as an additional performance measure the proportion of noncredit students who end up entering regular credit programs at community colleges. But data on non-credit students are not easy to aggregate at most community colleges because they are not usually included in regular student registration records but are, instead, kept in specially-constructed databases. Similarly, important academic experiences such as non-course-based academic skills-building interventions delivered through individualized tutorials or dedicated skills centers are not captured in regular student records systems because they are not "courses." Community colleges need to make it a priority to integrate these data systems so that students beginning in non-credit environments are included in tracking and fairly counted as potential contributors to key outcomes.

Remaining Challenges

While the last ten years has seen substantial progress in developing new data resources and measures for community colleges, there are a number of challenges that remain to be met. Among the most prominent are:

- Integration of data about credit and non-credit students and course-based versus non-course-based experiences, as above.

- Creation of state-level SUR databases in the states and systems that currently lack such capacity, as well as inclusion of non-public institutions (private and proprietary) to increase the number of potential transfer destinations (SURs in 19 states now include at least some of these institutions).

- Development of standard (and widely accepted) definitions for performance measures and descriptive variables about

students and academic experiences that are outside the scope of current federal reporting.

- Re-regulation or revision of the Family Education Right to Privacy Act (FERPA) to clarify the original purposes of the Act to enable research to improve postsecondary education to be undertaken more effectively, while preserving individual rights to privacy with appropriate safeguards. Many potential users of state SURs and registration records are deterred from tapping them for research purposes because of false but persistent perceptions of what FERPA does and does not allow.

- Development of standard calculation algorithms for these measures that can be made widely available to community colleges pre-programmed in off-the-shelf software environments like Microsoft Excel, Microsoft Access, SPSS, or SAS.

By far the most important challenge for community colleges in the data and information arena, though, is the fact that so much available data is not used. There are many reasons for this including lack of knowledge and awareness about how to link data to presenting problems, shortfalls in campus analytical and institutional research capacity, and lack of knowledge about how to present data to faculty, staff, and stakeholders in a manner that tells an action-related "story" about what is happening and what needs to be fixed. A concerted multi-year national effort on how to remedy these conditions, as is currently undertaken by such national initiatives as Achieving the Dream and Complete College America, would likely pay major dividends with respect to using data for improvement. The necessary tools have already been developed (see next section). But it will take sustained effort, perhaps kicked off by another White House Community College Summit focused explicitly on using data for improvement, to move the needle on this issue.

Resources

Fortunately, the last decade has seen considerable progress in developing appropriate resources to support community colleges and states in developing more powerful student databases and putting them to use. Among these are both publications and organizations that can provide technical assistance. Important publication resources include:

- *The Community College Data and Information Toolkit* (www.**communitycollege**central.org/.../**Data**_Performance_**TOOLKIT**.pdf). This publication by the Community College Bridges to Opportunity program at the University of Texas Austin contains all the basics needed to understand and use data in a community college setting.

- *Strong Foundations* (http://www.sheeo.org/sspds/default.htm). This is an up-to-date inventory of the contents, capacity, and applications of state Student Unit Record databases.

- *Using Longitudinal Data to Increase Community College Student Success* (http://ccrc.tc.columbia.edu/Publication.asp?UID=570). Prepared by researchers at the Community College Research Center at Teachers College, Columbia University, this is an accessible guide to creating longitudinal databases at individual colleges and using them to create a variety of performance and outcomes measures.

Important organizations that can provide assistance include:

- The Community College Research Center at Teachers College, Columbia University. Faculty and staff at this center have conducted innumerable studies and demonstration projects with foundation support and under contract (http://ccrc.tc.columbia.edu/). A similar resource is the Community College Leadership Program at the University of Texas Austin (http://edadmin.edb.utexas.edu/cclp/).

- The National Center for Higher Education Management Systems (http://www.nchems.org/), the State Higher Education Executive Officers

(http://www.sheeo.org/), and the Western Interstate Commission for Higher Education (http://www.wiche.edu/). These organizations specialize in providing guidance about how to best harness student records data to construct institutional research and policy studies to improve student success.

- Jobs for the Future (http://www.jff.org/). This multi-purpose applied research center provides assistance in designing and implementing data-driven student success programs directed primarily at underserved students.

- The National Community College Benchmark Project (http://www.nccbp.org/). Located at Johnson County Community College, this is a national membership consortium of community colleges that share commonly-defined data about institutional characteristics and effectiveness.

- The Diversity Scorecard Project (http://cue.usc.edu/equity_model/). Located at the University of Southern California, this initiative has developed innovative and effective approaches to packaging disaggregated community college performance data to induce action by faculty and staff.

Summary and Conclusion

Converting data into information and using this information to diagnose what impedes and accelerates student progression in community colleges is critical to achieving the national attainment goals set at 60% of young adults in the U.S. by President Obama. Good data systems, well utilized, enable college leaders, faculty, and staff to determine what works specifically for which kinds of students in the complex, diverse, and challenging environments provided by today's community colleges. Sound and well-utilized data systems are as important to student success as dedicated and well-prepared faculty, caring and knowledgeable student support staff, and up-to-date and appropriate educational facilities and technologies. We cannot succeed without them.

References

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

Finney, J. E and Stoel, C. F. (2010). Fostering Student Success: An Interview with Julie Phelps. *Change*, 42(4) July/August 2010, 38-43.

Garcia, T. I. and L'Orange, H. P. (2010). *Strong Foundations: The State of State Postsecondary Data Systems*. Boulder, CO: State Higher Education Executive Officers (SHEEO).