LUMINA®
FOUNDATION FOR EDUCATION
NEW AGENDA SERIES™

Critical Connections:

Linking states' unit record systems to track student progress

Peter Ewell and Marianne Boeke
National Center for Higher Education Management Systems

January 2007

Acknowledgments

We gratefully acknowledge the 50 state informants and the 47 state or system database administrators who gave generously of their time to provide us with documentation and answer our questions. Their cooperation was critical to the completion of this project. We also thank Lumina Foundation for Education for the support needed to carry it out.

The views expressed in this report are those of the authors and do not necessarily represent those of Lumina Foundation for Education, its officers or employees.

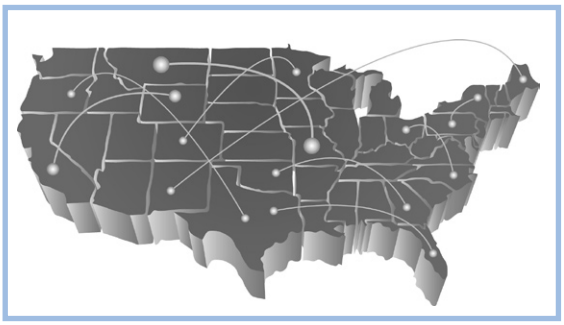
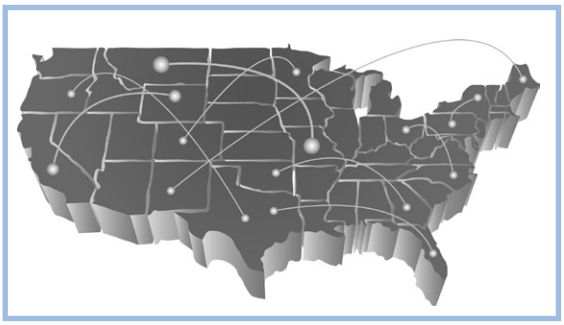


Table of contents



Background

In April 2003, Lumina Foundation for Education published a comprehensive inventory of state-level student unit record (SUR) capacity — an inventory conducted in 2002 by the National Center for Higher Education Management Systems (NCHEMS) and titled *Following the Mobile Student: Can We Develop the Capacity for a Comprehensive Database to Assess Student Progression?* (Ewell, Schild and Paulson, 2003). The primary intent of this effort was to examine the feasibility of linking student data from multiple states to track students on a national basis.

Since that 2002 survey, interest in SUR databases has grown considerably. The need for states to develop capable K-12 databases to meet the reporting requirements of the No Child Left Behind (NCLB) law has stimulated vigorous efforts to design and implement new systems; it also helped fuel the development of the national Data Quality Campaign (DQC) to promote best practice in this arena.[1] One of DQC's "ten essential elements" for state K-12 longitudinal data systems is the ability to link individual student records from secondary schools to higher education systems; the most straightforward route to doing so is to match K-12 records with postsecondary SURs in those states that possess them.

At the same time, the National Center for Education Statistics (NCES) proposed a national SUR system. Soon after, two prominent panels — the National Commission on Accountability (SHEEO 2005) and the Secretary of Education's Commission on the Future of Higher Education (USDOE 2006) — both recommended such a system be developed. These recommendations speak directly to the need to track student progress on a national basis, and such a system would give state SUR databases — which frequently are far richer in data than the proposed federal system — a mechanism for following students beyond state lines.[2]

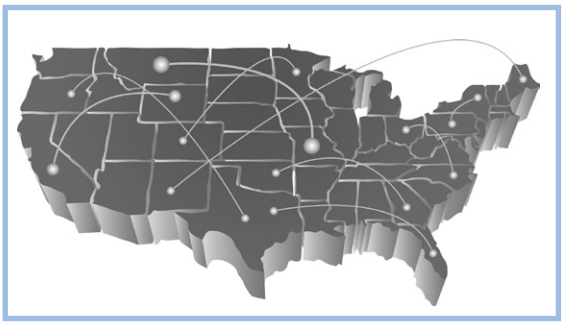
These recommendations speak directly to the need to track student progress on a national basis.

Meanwhile, state policymakers see a growing need for solid longitudinal information about student progression. Calls for increasing the flow through the "educational pipeline" are growing louder in state policy discussions as business and civic leaders recognize how vital this "supply chain" of educational capital is in their states (Miller and Ewell 2005). And the task of increasing student achievement levels is tougher than ever. The United States has fallen behind other nations in the proportion of young adults with a postsecondary credential, and the nation continues to show significant gaps in attainment between white students and students of color.

These changes — along with the rapidly developing sophistication of state SUR resources in higher education — show that it is time to re-examine "the state of the states" in this important area. This report presents the results of NCHEMS' new 50-state inventory of state SUR capacity. In addition to revisiting the topics covered in the 2002 survey (data-element coverage and analytical capability), the current inventory more closely examines how states link SUR data with other data sources and how they use the resulting information. It also examines some of the cross-cutting issues that states face in this arena.

The report is organized into six sections:

- The first section describes the **methodology** used by NCHEMS to conduct the 50-state survey and analyze its results.
- The second presents results on the **overall status** of state- and system-level SUR systems on a number of dimensions, including the proportion of states that have such databases and their overall capabilities.
- The third takes these findings to a finer level of detail by examining the **specific data elements** contained in these systems and the extent of cross-system commonality. It also examines the capacity of states to generate a "common core" of data elements suitable for consistent tracking.
- The fourth section examines the most **common applications** of state SUR data through regular reporting.
- The fifth section presents cross-cutting **challenges and issues**.
- The report's final section provides some **recommendations** for moving forward.



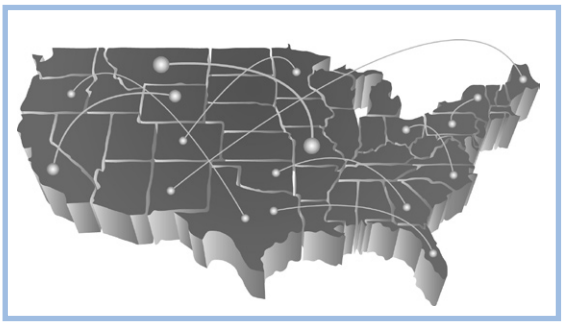
Methodology

The basic approach used to conduct the 2006 inventory was broadly similar to the method NCHEMS used in 2002. Staff first identified and updated the contacts used for each state in 2002 and sent them an initial letter describing the project and the type of information sought.[3] We then asked each contact by telephone or e-mail to provide extensive written documentation about their state's SUR database, including (if available) overall descriptions of the system, institutional reporting instructions and/or formats, data element definitions and dictionaries, applicable data structures and record layouts, and examples of the kinds of reports generated by the system. NCHEMS staff members then reviewed this primary source material to develop initial answers to questions contained in a standard protocol.[4] Gaps in topical coverage were addressed through follow-up phone or e-mail interviews — a process that frequently required several rounds of contacts.

As in the 2002 project, we used this approach rather than the simpler, more common method of sending a survey to state contacts. We did this because we believed that direct examination of source documents would reveal more about actual state database capabilities. Using this method, we successfully contacted all 50 states and obtained usable documentation for 47 databases in 40 states. When data collection was complete, we summarized the results in a write-up for each database, sending drafts of these write-ups to respondents for final verification.[5]

Survey results were compiled and tabulated in Microsoft Excel. In addition, as in the 2002 survey, we used a standard data element template based on NCHEMS' extensive work with states and institutions in the development of longitudinal data systems (Ewell, Parker and Jones, 1988). For each data element on this list, the template addressed such matters as coding structures and source of definitions.

The resulting Excel spreadsheets were used to conduct a series of descriptive analyses addressing such topics as the number and percentage of states maintaining SUR databases with particular capacities, and the proportion of states maintaining different combinations of "common core" data elements.



Overall status of state SUR data systems

Forty of the 50 states have operational SUR databases covering public institutions in their states (see Table 1, Page 25). This represents a gain of one state (Kansas) since the 2002 survey and a gain of seven since 1999 (Russell, 1999). The ten states not covered by SUR databases are, for the most part, fairly small. Among them, only Pennsylvania and Michigan have substantial numbers of higher education institutions and enrollments. As a result, the actual coverage of these databases is greater than it appears with respect to enrollments, though not for institutions. Eighty-one percent of the nation's total headcount enrollment and 77 percent of its FTE enrollment is collectively covered by current state SUR systems.

These database systems share a number of important characteristics:

- **Multiple databases**. Seven of the 40 states that maintain SURs do so through more than one database. These states contain several independently governed public higher education systems, each of which maintains its own student information for planning and reporting. One state (California) maintains three SUR systems governed by the three sectors of higher education in that state.[6] Others maintain separate SUR systems for the two-year and four-year public sectors (North Carolina, Oregon, Washington and Wyoming), and one (New York) maintains two databases corresponding to the two public systems of higher education (CUNY and SUNY). In most cases, states with multiple databases have established mechanisms to share unit record data with one another. But in some cases, a state's four-year sector supplies only aggregate reports to the community college system.

- **Institutional coverage**. Most SUR databases include information only on public institutions within the state. Although all public enrollments are covered by the SUR databases of the 40 states that maintain them, 17 states now include at least some information drawn from independent institutions — a gain of five states since 2002. Four of these (Florida,

Oklahoma, South Carolina and Texas) include data on all independent, nonprofit institutions in the state, with three additional states (Colorado, Minnesota and Virginia) about to join them with full independent participation. Most of the remaining states in which independent colleges participate (Arkansas, Illinois, Kentucky, Maryland, Massachusetts, New Jersey, New Mexico, New York and Ohio) include records on most independent institutions. Two others (Illinois and North Carolina) include records from just a handful of independent institutions. Six states include records from at least some proprietary institutions.

Overall, the trend among non-public institutions is toward greater participation in SUR databases. Respondents in nine states report that they plan to increase the institutional coverage of their databases. While most independent institutions that participate in state SURs do so mainly to take part in state scholarship programs, more independent colleges are expressing interest in obtaining and analyzing this data. Proprietary institutions are typically omitted from state SURs, although Virginia plans to include all proprietary institutions in its upcoming data cycle, and Florida has greatly expanded proprietary participation.

However, even though independent and proprietary institutions are sometimes included in SUR data collection, the full range of data is often unavailable. For example, several states include only those students enrolled at private colleges who participate in state scholarship programs. Others only collect data for independent attendees who are state residents. Similarly, most states collect fewer data elements for students attending private institutions and do so less frequently. For example, many states only collect fall term enrollment information and degrees granted for independent institutions.

Overall, the trend among non-public institutions is toward greater participation in SUR databases.

- **Historical data**. Most state SUR databases have been in place long enough to support meaningful longitudinal analyses of student progress and degree completion. Nine of the 47 databases date back to the 1970s, and 14 were established in the 1980s. Of the remainder, only eight are less than 10 years old. When an SUR system is created, two years of formative work are typically required before meaningful data can be collected. That means that all but these last eight databases are robust enough to allow researchers to track several longitudinal cohorts over the federally established Graduation Rate Survey (GRS) period of six years for four-year institutions.

- **Data detail and periodicity**. All of the SUR databases examined contain the standard student descriptors needed to complete federal IPEDS reports, including gender, race/ethnicity, enrollment, degrees granted and program major.[7] All but eight of 47 contain annualized information about credits completed, and about half of them track grades or grade-point averages. Twenty-three collect data at the transcript level — that is, a record is maintained for every student enrolled in every course, and a growing number of states express interest in moving to this level of detail. By far,

the most data in these systems concerns credit-bearing work, although 12 of the systems examined include at least some data about non-credit activity.[8] All but five of the databases examined contain at least some data on financial aid, and 25 include considerable detail on that topic.

With regard to periodicity, all but two SUR databases capture enrollment data on a term-by-term basis, although most gather degree-completion data annually. In some cases, these data are gathered retrospectively — that is, there is only one annual data capture for institutions, but it includes term-level detail for the prior year. All but 18 employ at least two capture points per term ("census date" and end of term), and a few have many more.

About half of the states with SUR databases have at least some experience in linking unit record data.

- **Record identification**. All SUR databases require a unique identifier to match student records from term to term or to add new information to an established record. Historically, this identifier has been the Social Security number (SSN), but this is beginning to change for a number of reasons. First, concerns about identity theft and a number of high-profile examples of on-campus security breaches have led more institutions to abandon the SSN as the primary identifier in their registration and student records systems. Second, new SURs that are being developed by all states under NCLB for elementary/secondary education feature unique identifiers not related to the SSN; a number of SURs in postsecondary education plan to adopt these identifiers at some point. But states must maintain student SSNs to link student records to external databases such as employment and military records. Also, institutions must continue to collect them so they can report student tuition payments to the Internal Revenue Service (IRS).

  In 2002, virtually all of the state SUR databases examined used the SSN as a key link. In the current inventory, 11 of the 47 databases examined have stopped using the SSN as the major key link; 12 more encrypt the SSN before manipulating the data. Two states report that they plan to stop using the SSN as a key link in 2007, and 10 more report eventual plans to discontinue its use. Despite the growing visibility of state efforts to create a new student identifier in K-12 education, only four states explicitly mentioned that they were planning to use (or even carry) this number in their postsecondary databases.[9]

- **Linking to other databases**. About half of the states with SUR databases have at least some experience in linking unit record data with external databases such as high school and employment data. Twenty-three of the databases examined report links with state unemployment insurance (UI) wage records to determine field of employment and earnings, and most of them do so regularly through a match on the SSN. Three additional states report plans to do this.

  Most of the states that make such linkages rely on the agency responsible for the UI records to do the match, but several have brought the matching function back into the state higher education agency. They have made this move because of

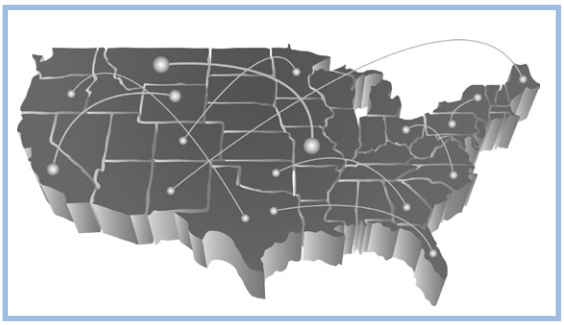
the perceived need to control student records more closely under the federal Family Educational Rights and Privacy Act (FERPA). Meanwhile, despite the growing salience of state-level longitudinal databases in K-12, only 11 states have linked college student data with high school records.

This will likely change quickly, though, as 16 additional states report that planning for such linkages is under way. Nineteen states have at least occasionally linked student records in higher education with other state unit record databases — including military records, federal employment records, driver's license records, and state records on public assistance and incarceration – though only a handful have tapped more than one of these. Finally, only five states have joined with neighboring states to share data about students' interstate mobility. Florida, Missouri, Kentucky, Ohio and Washington have some experience with this, and a four-state data-exchange demonstration coordinated by NCHEMS is currently under way.

Taken together, these results show little change since 2002 in the number of states with SUR data resources in higher education. Still, the results do indicate significant changes in institutional coverage and sophistication. The growing number of states that include all independent institutions and those that now include proprietary institutions is especially striking in light of private college associations' vigorous opposition to establishing a federal unit record system.

> The only **substantial** area in which little **progress** has been made is in **sharing** across state lines.

At the same time, many states are enhancing their SUR data systems by adding new data elements or additional data capture points. And even though FERPA constrains such efforts, a significant number of states are finding ways to link their data with external databases. The only substantial area in which little progress has yet been made is in sharing data across state lines.



Specific data elements, definitions and coverage

As in 2002, the new NCHEMS inventory allowed direct investigation of the content and structure of the 47 state-level SURs identified because state data personnel were asked to provide full documentation of their systems. This enabled us to examine each database against a group of 33 data elements — a set that is sufficient to meet most longitudinal tracking and reporting needs (Paulson, 2002). We first examined the documentation for each database to determine whether each of these data elements was present. If a given data element was present, we sought further information about how it was coded and defined. We then looked for important combinations of data elements that would have to be simultaneously present and consistently defined to allow for certain kinds of analyses.

Basic data element coverage. Table 2 (Page 26) provides a breakdown of overall data element coverage by reporting the number and percentage of SUR databases (N=47) that contain one or more of the 33 key data elements, together with corresponding statistics for all states (N=50) and states with SUR databases (N=40).[10] Data elements are reported as present in this analysis under the strictest definition possible. That is, if there was any doubt about whether a particular data element was present in a particular database, that element was counted as not present. This means that the statistics presented in Table 2 represent minimum estimates of coverage for each of the 33 key data elements.

A number of conclusions are apparent from this display. First, all current state-level SURs contain sufficient information to track student enrollment on an annual basis through degree completion. Further, this annual tracking can be broken down by a number of important demographic variables, by full-time/part-time status and by program of study. All but two of the state-level databases can do this on a more detailed term-to-term basis; this can help officials better monitor patterns of student "stop-out" behavior and in-state transfer. Looking at academic performance longitudinally, databases in 33 states can report on accumulated credits and

grade performance, with 22 of these being able to report credits and 19 able to report grades on a term-by-term basis.

Second, most state-level SURs provide some information about student academic background. Databases in 25 states can generate admissions test scores; those in 34 states can identify the high school a student attended, and those in 38 states can pinpoint when a student graduated from high school. Beyond these elements, however, data on academic background is more limited. Only about half of these databases contain data on high school performance, with about three-quarters recording prior colleges attended. Only about a third contain detailed data about student remediation status or placement test scores. Still, the proportion of state-level SURs that do maintain these academic background data elements has increased notably since the 2002 inventory.

Most state-level SURs provide some information about student academic background.

Third, there has been substantial progress in the number of state-level SURs collecting other kinds of data important for planning. Reflecting considerable growth in these areas, more than half now include data on high school concurrent enrollment. Also, more than half include data on student participation in distance or mediated instruction (up from only about 15 percent in 2002). Almost 90 percent include at least some information on student financial aid, with just over half containing considerable detail about financial aid.

Coverage for combinations of data elements. A viable assessment of SUR data capacity goes beyond the presence or absence of particular data elements. Another salient point is the extent to which designated combinations of data elements exist – combinations that can enable specific types of reports to be generated. Table 3 (Page 27) provides results for five important combinations of data elements, with coverage statistics calculated in the same manner as those in Table 2. These combinations are organized in order of expanding inclusiveness of data elements, beginning with the least common denominator across states. Each successive combination adds some new elements to the preceding cluster, and three clusters are additionally disaggregated by whether or not they include data on student grades.

Table 4 (Page 28) displays results broken down by individual database, presented on a state-by-state basis. An "X" in a given column in Table 4 indicates that all members of the particular data element cluster are present. Once again, it is important to note that the most conservative approach was applied to determining whether a particular data element was present in a given database, so the statistics presented represent minimum estimates of the actual coverage.

The "core" combination consists of gender, race/ethnicity, date of birth, current enrollment, program/major and degree awarded. This combination of data elements allows calculation of annual student progression and degree-completion rates across institutions by academic program and for a range of important demographic subpopulations. All existing SUR databases contain this combination, representing 40 states. Coverage for additional clusters of data elements is as follows:

- Set 1 adds geographic origin and full-time/part-time status to the core. This allows for additional disaggregation by region and additional reporting distinctions based on pattern of enrollment. This cluster exists in all but three SUR databases in all but two states.

- Set 2 adds basic term-level performance information in the form of credits attempted and credits earned. This allows estimates of course-completion rates, even in the absence of transcript-level detail, as well as overall credit-completion ratios. It also allows for more precise breakdowns of persistence and completion rates by enrollment intensity. Having term-to-term statistics on these additional data elements also means that officials can examine patterns of stop-out behavior for different student populations. (It's important to note, however, that only annualized statistics can be used to compare performance because different states and systems operate on different term structures; e.g., semesters and quarters.) Set 2a adds grade information to this mix. Four in five databases contain the data elements in these sets, representing about 60 percent of the states.

- Set 3 adds high school information to Set 2. This enables production of detailed feedback reports to high schools on student persistence and course completion in college. Set 3a adds grade information, allowing such performance reports to include GPA. About two-thirds of current SURs in about half of the states contain this particular combination.

- Finally, Set 4 adds financial aid information to Set 3. This enables officials to study the effectiveness of aid investments in promoting the success of various types of students. Set 4a includes grade information, allowing such reports to include GPA. Slightly fewer than half of current SURs in about a third of the states contain this particular combination.

Data element definitions. The range of data element coverage reported in Table 2 provides a solid foundation for most states with SUR databases to undertake a range of detailed longitudinal studies of student progression within a state's boundaries. Comparative analysis or multistate data-sharing efforts, however, will depend a great deal on the consistency with which data elements are defined across states. Accordingly, Table 5 (Page 30) presents a definitional review of the 33 data elements contained in Table 2.

> Much will depend on the consistency with which data elements are defined across states.

The first column in Table 5 classifies the current state of definitional consistency across SURs into three levels. The second column provides the basis for this classification. The highest level of commonality is assigned to a given data element if a) definitions and code structures are identical across the databases examined or, b) if definitions are identical and code structures are sufficiently compatible that an identical entry could be constructed through concatenation or recode. A moderate level is assigned if there is basic definitional consistency across databases but a) different code sets are used by different SURs or, b) different but compatible objects are referenced by the same data element. In such cases, definitional consistency can be achieved in principle by establishing an equivalency table for each code set or a list of the entities in use. Finally, a low level is assigned if substantially different definitions are used in different SURs or, as is more frequently the case, if the actual definitions applied are unknown or institutionally maintained.

A review of Table 5 suggests that there is basic compatibility across state SURs for most of the

33 data elements examined, but that some effort would be required to create complete equivalency. Also, the degree of compatibility differs for different kinds of data elements.

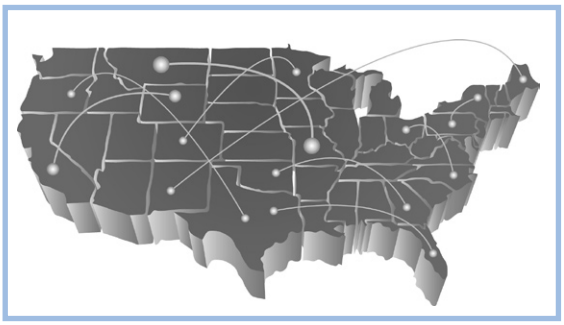
- **Demographics**. Definitional commonality is highest for this set of data elements. For each element there is either an IPEDS definition in place or states use the same basic definition, while sometimes employing a slightly different coding structure. For example, date of birth is most commonly presented in DD/MM/YY format, but some SURs use a reverse order or only record a student's year of birth. Similarly, most systems record home county in geographic origin for in-state students and record state or country of origin for out-of-state students. Still, some systems only record state of origin. These minor inconsistencies can all be addressed through concatenation or recoding.

- **Academic background**. Definitional commonality is moderate for this set of data elements, for a variety of reasons. Some data elements — for example, high school attended — refer to the same entity but use entirely different code sets (e.g., IPEDS Unit ID vs. CEEB code). Others — for example, admissions and placement test scores — refer to similar but different entities (e.g. SAT vs. ACT scores or as many as a dozen different placement tests). Achieving full definitional consistency under these circumstances would require the establishment of equivalency tables that include all entities and code sets in use.

- **Enrollment status**. For basic enrollment factors such as degree-seeking status and full-time/part-time attendance, there is a high degree of compatibility across data elements. For enrollment situations such as concurrent enrollment or distance education, on the other hand, there are no established national definitions or conventions, even though coding structures may be identical across databases. Finally, financial-aid information is kept at widely varying degrees of detail in different SURs. Equivalency tables, as above, could handle the majority of these situations, but some would require the establishment and application of common national definitions.

- **Academic activity and attainment**. The main challenge to compatibility among the data elements in this set is the fact that different states and systems operate on different term structures. Although most use standard 16-week semesters, some use quarters, and virtually all differ somewhat in the way they handle summer terms. Consequently, unless aggregated to yield annual totals, these entries are incompatible. At the aggregate annual level, however, they can be made equivalent so long as the credit values are adjusted. For SURs that maintain cumulative credits earned and GPAs, these entries are compatible across databases so long as similar adjustments are made in credit and grade calculations. Degree awards vary in the level of detail carried but can be aggregated to standard IPEDS categories.

Overall, it is clear that working equivalencies can be established to enable the vast majority of data elements to be compared across state contexts and aggregated to yield multistate totals. The principal challenges involved in widening comparative efforts lie in institutional coverage and data element contents.



Common applications in data use and reporting

States and systems use the data contained in SUR systems for a wide variety of reporting, although financial constraints prevent many of them from using the data as much as they would like. Most of these reports are publicly available via the Internet, either on the SHEEO Web site or on individual systems' Web pages. Among the most common regular reports and analyses undertaken using SUR data resources are the following:

- **Graduation and retention**. By far, states' and systems' most common use of SUR data resources is to calculate graduation rates. Forty-four of the 47 states or systems with SUR databases report graduation rates, although two of these do not use data from their SURs to do so.[11] All use a method consistent with the federal Graduation Rate Survey (GRS) to calculate six-year rates for four-year institutions and three-year rates for two-year institutions. Twenty-one use the GRS method essentially without alteration, but many of these allow additional years to complete and/or include part-time students as a separate cohort. Eighteen more calculate a parallel set of completion statistics based on completion at any institution within the state. Forty-one calculate first-year retention according to GRS conventions, and 18 (the same 18 that report in-state completion) list students as retained if they are enrolled at any institution in the state. All of these statistics are reported for individual institutions. Only three disaggregate these data further by gender and race/ethnicity.

- **Transfer and multiple enrollments**. Thirty-four of the states or systems with SUR databases calculate and report transfer rates from public two-year colleges to public four-year colleges, although only about a third of these report on academic performance after transfer. Eight examine wider transfer patterns among all institutions in the state for which they have data. Six have arranged to obtain additional

transfer information from the National Student Clearinghouse.

- **Job placement/workforce development**. Seventeen states or systems generate regular reports on student employment and/or earnings after program completion. Most of these are directed at employment within the field for which a degree or certificate was granted, and all but two are based on record matching with UI wage record files.[12] Five additional states have undertaken more in-depth studies of workforce development, including the role of non-credit vocational training.

- **High school feedback**. Ten of the states or systems with SUR databases construct detailed high school feedback reports. These reports provide a number of indicators of postsecondary progress for former students in each of the state's high school districts. The most common indicators used in these reports are students needing remediation (by field), credits attempted and earned, grade-point average and first-year persistence. One additional state produces a feedback report that only provides the proportion of students needing remediation. Two states regularly monitor and report on high school concurrent enrollment for college students.

- **Developmental/remedial education**. Ten states or systems generate regular reports on the proportion of students at each institution in need of remediation (for the most part by field). Seven of these track remedial students into college-level work and report on the results. Respondents to our inventory also reported several recent in-depth studies of remediation. These studies examined such matters as the effectiveness of current placement policies, appropriate cut scores on different placement tests, and the relative effectiveness of various ways to deliver developmental education.

- **Distance education**. Finally, eight states or systems generate regular statistics on the extent of participation in distance education programs. However, definitions of what constitutes "participation" range from enrollment in a complete online degree program to simply taking an online course in a given academic year.

Beyond the particular topics addressed, states and systems engage in a number of common reporting practices. Among the most prominent are:

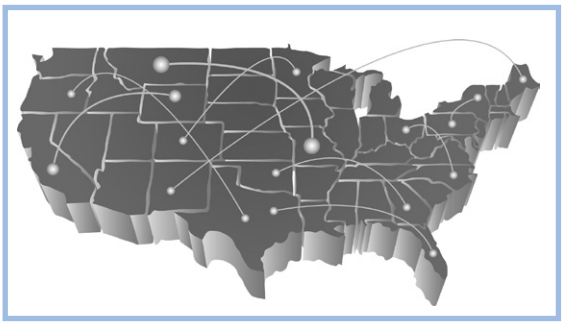
- **Performance indicators**. Twenty-five states or systems use SUR data to construct one or more performance indicators to track progress and discharge accountability. The most common are graduation and retention rate statistics, but many systems include course-completion rates, job placement and remediation rates, and success in remediation. About 10 indicators are usually calculated, but some states compile many more. Performance indicators are generally posted prominently on the agency Web site. Most states or systems display institutions' performance on each indicator side by side to facilitate comparison across institutions. Two states, however, only compute performance indicators for the system as a whole.

States and systems engage in a number of common reporting practices.

- **Institutional profiles**. Although all 47 of the states and systems examined compile basic statistics on enrollments for public institutions, 13 of them compile readily accessible "institutional profiles." Each profile provides about a dozen key statistics about institutional enrollments, programs, retention/completion and other characteristics. These profiles are frequently presented graphically and appear consciously intended as "consumer guides" for potential students and interested stakeholders.
- **User access**. Finally, six states or systems include a utility on their Web sites that enables users to actively manipulate SUR data to generate customized reports about enrollments and other topics. Only two of these allow direct access to unit records themselves; the other four allow use of pre-built tables or data cubes that can be manipulated to create customized reports. In the two cases where direct access is possible, identifiers have been stripped from the record, and the report generator employs a minimum cell-size convention to protect confidentiality. Despite the relative infrequency of direct user access at this point, states are clearly moving in this direction; a number of respondents said they were planning to develop such a capability.

> Most states and systems say that about **half** of their time using the SUR **database** is in response to **ad hoc** queries.

In addition to regularly produced standard reports, most states and systems say that about half of their time using the SUR database is in response to ad hoc queries. Not surprisingly, most of these questions are posed by legislators, the governor's office or board members. Less frequently, the data are tapped to support a research effort centered on a particular issue — for example, the efficacy of teacher-training or workforce-education programs.



Cross-cutting issues and challenges

Finally, respondents to the inventory reported a number of common issues and challenges in the establishment and operations of SUR data systems. Among the most prominent of these were:

- **Data quality**. In light of the NCES proposal to create a national student unit record system for higher education, it is striking how many states or systems mentioned the effort involved in obtaining consistent and reliable data from institutions. More than half of respondents noted this problem. Many reported that institutional responsiveness was uneven and that obtaining usable unit record data required substantial diplomacy and a good deal of "hand holding." Part of the problem is that staff turnover at agencies and institutions demands nearly constant retraining. At several agencies, a decision has been made to decrease the frequency of reporting with the hope of getting better-quality data. Whatever the approach, moreover, states and systems invest considerably in cleaning and verifying data — processes that are costly in staff time and extend the time needed to prepare reports and analyses.

- **Personnel shortages**. A substantial majority of respondents report that the second-biggest challenge they face with SURs is understaffing, particularly among the critical programming staff who keep these databases operational. Staff turnover in such positions is high, and state salaries are not competitive. In addition, budget cuts in many states have fallen disproportionately on technical positions because these positions are deemed less essential than operational positions. As a result, systems are inadequately maintained and, in some cases, data collection and updating have been temporarily suspended. As one respondent put it: "I would like to say we are just barely keeping our heads above water, but I am not sure we have our heads above water."

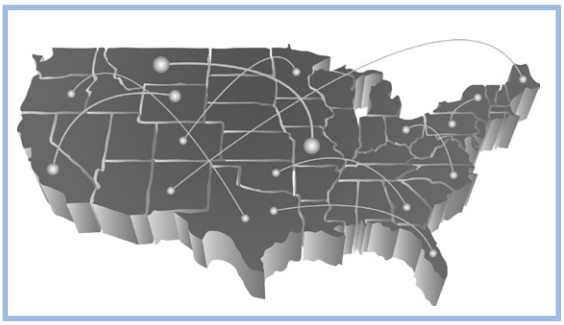
- **Reporting and analysis**. In addition to these more general personnel challenges, most respondents noted that there was

little time or staff available to perform analyses using their increasingly rich stores of SUR data. The following comment was typical: "There is more we could do if we had the resources … currently we limp along, knowing what is possible and what we could do, but simply can't at this point." Others noted that they have only a few staff members dedicated to analysis and therefore few colleagues or opportunities for professional exchange and development. As a result, they often undertake analyses on their own with few examples to follow. Finally, many respondents noted the challenges associated with educating their "customers" — the many governmental and public stakeholders who are looking for information on higher education. Frequently, such requests reflect significant misunderstandings of what particular statistics mean or their range of applicability.

- **FERPA and privacy issues**. Issues associated with FERPA and records confidentiality have become much more salient in the past four years. Although few actual security breaches have occurred in higher education (and none in SHEEO agencies), the "perception of FERPA" has clearly had a chilling effect on state and system use of SUR data. This has been particularly true with regard to record matching — either across higher education databases or between them and K-12 or UI wage records. Linking individually identified records outside the direct control of the agency is increasingly being cited as a form of "redisclosure" which is prohibited by FERPA. This has led a number of states to bring all data-matching procedures in house, where the link can be accomplished by agency personnel under tightly controlled conditions. Unfortunately, the more common result has been that states and systems simply stop doing valuable things to track students' progress.

> The "perception of FERPA" has clearly had a chilling effect on state and system use of SUR data.

It is interesting to note that technical issues were not on most respondents' list of significant challenges. A few states did report that they were in the middle of major system conversions that were consuming substantial amounts of staff time, and a few more noted that they were evolving to more capable and modern database environments. Still, reports of these types of technical challenges were far outnumbered by comments centered on the four issues noted above.



Implications and recommendations

Results of the current inventory suggest a number of conclusions and next steps for the development and use of state SUR data resources in higher education. Some of these conclusions and recommendations build on the main finding of the 2002 study: that it is feasible to further harness these databases to create a national capability to track student progress. Others have been bolstered by other, more recent trends, including the recommendations of the Secretary of Education's Commission on the Future of Higher Education, the proposal to create a federal student unit record system, and the growing use of SUR databases in elementary and secondary education. In light of these developments, we recommend the following action steps:

- **Establish a common core of data elements**. This inventory suggests that the nation is already quite close to establishing a consistent "common core" of data elements to populate SUR systems. In some cases, standard definitions of these data elements are in place through IPEDS or the Common Dataset, but in the majority of cases such definitions vary somewhat from state to state. If these substantial data resources are to be more useful on a national basis, efforts must be made to standardize a set of key data elements across states and systems. The 33 data elements analyzed in this study represent a good starting point for this "common core," but cross-state action and deliberation are needed to determine whether additional data should be included.

  Once the contents of a common core are determined, there might be two ways to proceed. The most straightforward would be to establish a common set of definitions and codes. This method would be difficult to implement, however, because reporting conventions in many states would have to change. Probably the more practical approach would be to develop an inclusive set of crosswalk tables, concatenations and recodes to convert the various current approaches to a compatible standard.

- **Identify and share "best practices" in data analysis and reporting**. Given their

growing importance, extent and coverage, state SUR resources are underexploited as strategic information resources. This is largely due to the following two factors: First, increasingly stretched state agency staffs simply lack the time and people to undertake more than simple accountability reporting. Second, because they are professionally and geographically isolated from one another, state-based data analysts lack good models for the kinds of reports and practices that can successfully influence policy.

On the one hand, this suggests a common research and development effort to identify and share best practices with respect to SUR database analysis and reporting. A subsequent step, once such reporting conventions are established, might be common benchmarking and more standardized analysis of subpopulations. At the same time, this suggests more and better professional development for state agency data analysts. Such training would help analysts to better share promising practices and would encourage exchange and experimentation. Finally, recent efforts show that often-overwhelmed state agency staffs can be usefully supplemented to expand analytical capacity. This has proven true in Florida, for example, where academic researchers have been allowed access to SUR data to investigate questions of value to state policy as well as academic interest, with appropriate privacy and security safeguards in place.

Staffs simply lack the time and people to undertake more than simple accountability reporting.

- **Create a "third-party" data-matching utility**. Although many states and systems have begun efforts to link their SUR data with other state-level databases such as UI wage records and high school records, doing so remains a challenge. And once again, states have been forced to undertake these efforts on their own and in isolation. Meanwhile, growing legal questions about redisclosure associated with FERPA further cloud these efforts, leading some states to simply abandon data linking because it is too troublesome or costly.

  In light of this situation, it would be extremely beneficial to create a national capacity to link unit record data quickly and securely. Technology already exists to accomplish this task, and the efficiency of creating a single utility to handle the needs of many potential users is apparent.

  Although some could argue that the eventual establishment of a federal unit record capability would obviate the need for such an entity, it will be years before one can be created. Moreover, even if a federal unit record system existed, states still would need linking capacity in order to tap their own and one another's data resources. Furthermore, such a utility would be most effective if managed as a service by a third-party nongovernmental organization, either established by an interstate compact or as an independent nonprofit agency.

- **Better align and link K-12 and higher education data resources**. Because of the requirements of NCLB, most states are quickly developing SURs for their elementary and secondary education systems. But

so far in most states, there has been little coordination between these efforts and existing higher education SUR resources.

Linking K-12 longitudinal data to higher education is one of the "ten essential elements" for data systems being promoted by the national Data Quality Campaign. But, beyond this general directive, little detailed development work has been done to determine how this should be accomplished, what data elements should be defined in common, and how the actual linkage should be made.

Of particular importance here is the establishment of consistent unique identifiers. All states developing K-12 SURs are creating their own unique identifiers without much coordination with one another or with higher education authorities. Several state SURs in higher education report that they plan to carry these numbers once they are developed, but colleges and universities draw students from many states, so compatibility and consistency in these identifiers is a must.

Meanwhile, most state and system SUR administrators would like to move away from the SSN as a unique identifier, and some are already creating alternatives independent of those being used in elementary and secondary education. This points to a compelling need for a national conversation about aligning and linking K-12 and postsecondary data resources, perhaps through the DQC.

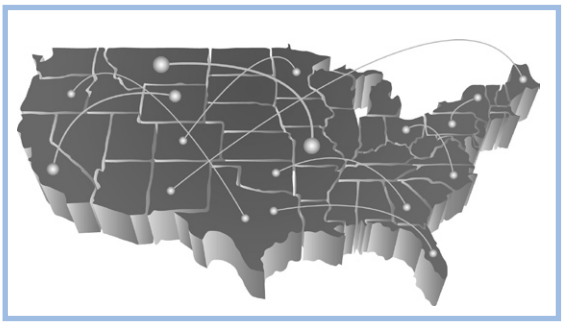
- **Clarify and re-regulate FERPA**. The Family Educational Rights and Privacy Act (FERPA) was never intended to impede legitimate educational research. Indeed, the law specifically exempts efforts by authorized agencies to undertake research designed to improve the effectiveness of education. Such work increasingly requires researchers to build longitudinal data files from multiple sources.

  State higher education agencies have a long history as ethical and responsible custodians of such records. Yet they are increasingly restricted by inconsistent interpretations of the law based on the principle that inaction is the safest course to follow. Particularly troublesome is the current interpretation of "redisclosure" to mean any attempt to use uniquely identified educational records in conjunction with similar records held by another responsible entity (state agency or institution). Such a prohibition exists nowhere else in state or federal operations, and it would be unheard of in the private sector, where record matching of personally identified data is routine.

  The DQC has noted this as an obstacle to developing state information capacity, though some progress is being made. Still, there should be a concerted, multistate effort on the part of higher education agencies to press the Department of Education to re-regulate FERPA so that the law reflects state and federal interests in improving evidence-based educational practices.

In sum, state-level SUR data resources in higher education are becoming steadily more inclusive and sophisticated. They now represent a considerable national asset that can and should be exploited more fully.

The existence and use of such data resources are a necessary condition for achieving the widely held policy goal of increasing the numbers of citizens who make it through the "educational pipeline" and attain postsecondary credentials. And these data resources are now at a point where a concerted, cross-state development effort would be a very wise investment.



Endnotes

1. NCHEMS is a Founding Partner of the Data Quality Campaign (see www.dataqualitycampaign.org).

2. This is a function that some states discharge partially through the National Student Clearinghouse.

3. The initial contact letter (Appendix A) is available, along with all of the supplemental material associated with this report, on the NCHEMS Web site: www.nchems.org/C2SP/SUR.

4. The e-mail interview protocol (Appendix B) is available, along with all of the supplemental material associated with this report, on the NCHEMS Web site: www.nchems.org/C2SP/SUR.

5. Each of the 47 individual database summaries (Appendix C) is available, along with all of the supplemental material associated with this report, on the NCHEMS Web site: www.nchems.org/C2SP/SUR.

6. Negotiations are under way in California to create a single SUR database using records drawn from the three system-level databases.

7. The fourth section of this report (Page 8) provides a more in-depth look at data element coverage.

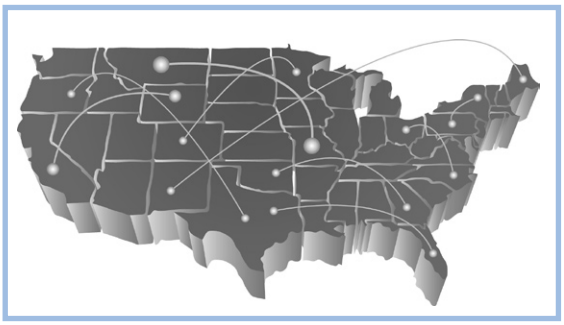
8. These data elements were chiefly concerned with ESL, ABE and GED activity.

9. The survey did not explicitly ask this question, however.

10. For states with multiple databases, the state was counted as maintaining a given data element only if all databases within that state maintain that data element.

11. Instead, institutions are asked to submit their locally calculated rates consistent with the GRS.

12. Two states use alumni follow-up surveys to gather this information.



References

Ewell, P. T.; Schild, P. R.; and Paulson, K. (2003). *Following the Mobile Student: Can We Develop the Capacity for a Comprehensive Database to Assess Student Progression?* Indianapolis: Lumina Foundation for Education.

Ewell, P. T.; Parker, R. J.; and Jones, D. P. (1988). *Establishing a Longitudinal Student Tracking System: An Implementation Handbook*. Boulder: National Center for Higher Education Management Systems (NCHEMS).

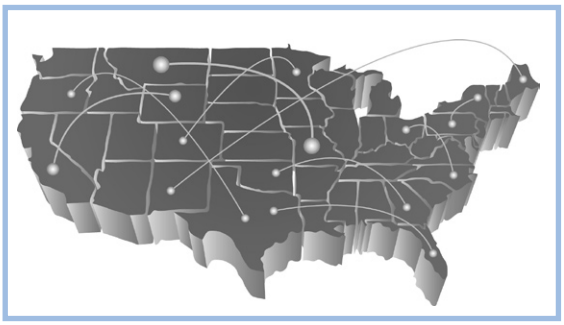
Miller, M. A. and Ewell, P. T. (2005). *Measuring Up on College-Level Learning*. San Jose: National Center for Public Policy in Higher Education.

National Commission on Accountability in Higher Education (2005). *Accountability for Better Results: A National Imperative for Higher Education*. Boulder: State Higher Education Executive Officers (SHEEO).

Paulson, K. (2002). *A Data Audit and Analysis Toolkit to Support Assessment of the First College Year*. Boulder: National Center for Higher Education Management Systems (NCHEMS).

Russell, A. B. (1999). *The Status of Statewide Student Transition Data Systems: A Survey of SHEEO Agencies*. Denver: State Higher Education Executive Officers (SHEEO).

The Secretary of Education's Commission on the Future of Higher Education (2006). *A Test of Leadership: Charting the Future of U.S. Higher Education*. Washington, D.C.: U.S. Department of Education.



Tables

Editor's note: Additional information from the authors — including detailed appendices and other state-specific supplementary material — is available on the NCHEMS Web site at: www.nchems.org/C2SP/SUR/.

Table 1: Student Unit Record (SUR) databases by state in 2006

State	Agency responsible	Date SUR established
Alabama	SHEEO agency	1998
Alaska	State university system	1997
Arizona	SHEEO agency	1998
Arkansas	SHEEO agency	1993
California	Community college system	1992
	University of California system	1980
	State university system	1970
Colorado	SHEEO agency	1987
Connecticut	SHEEO agency	1988
Delaware	[None]	
Florida	Statewide K-20 authority	1985
Georgia	SHEEO agency	1983
Hawaii	SHEEO agency	1995
Idaho	[none]	
Illinois	SHEEO agency	1983
Indiana	SHEEO agency	1979
Iowa	[None]	
Kansas	SHEEO agency	2002
Kentucky	SHEEO agency	1980
Louisiana	SHEEO agency	1977
Maine	SHEEO agency	[Unknown]
Maryland	SHEEO agency	1977
Massachusetts	SHEEO agency	1985
Michigan	[None]	
Minnesota	SHEEO agency	1983
Mississippi	SHEEO agency	1984
Missouri	SHEEO agency	1988
Montana	[None]	
Nebraska	[None]	
Nevada	SHEEO agency	2000
New Hampshire	[None]	
New Jersey	SHEEO agency	1985
New Mexico	SHEEO agency	1994
New York	State university system (SUNY)	1988
	City university system (CUNY)	[Unknown]
North Carolina	Community college system	1980
	State university system	1978
North Dakota	SHEEO agency	[Unknown]
Ohio	SHEEO agency	1998

Table 1, continued

State	Agency responsible	Date SUR established
Oklahoma	SHEEO agency	1977
Oregon	Community college system	1995
	State university system	1990
Pennsylvania	[None]	
Rhode Island	[None]	
South Carolina	SHEEO agency	1993
South Dakota	SHEEO agency	1998
Tennessee	SHEEO agency	1991
Texas	SHEEO agency	1973
Utah	SHEEO agency	1999
Vermont	[None]	
Virginia	SHEEO agency	1992
Washington	State university system	2000
	Community and technical college system	1994
West Virginia	SHEEO agency	[Unknown]
Wisconsin	State university system	1973
Wyoming	University of Wyoming	[Unknown]
	Community college system	2000

Table 2: SUR database coverage of key data elements

Data element	SUR databases		All states		States with databases	
	N	%	N	%	N	%
Demographics						
Sex	47	100.0%	40	80.0%	40	100.0%
Race/ethnicity	47	100.0%	40	80.0%	40	100.0%
Date of birth	47	100.0%	40	80.0%	40	100.0%
Citizenship	46	97.9%	39	78.0%	39	97.5%
Geographic origin	46	97.9%	39	78.0%	39	97.5%
Disability status	17	36.2%	11	22.0%	11	27.5%
Academic background						
Admissions test scores	31	66.0%	25	50.0%	25	62.5%
High school attended	38	80.9%	34	68.0%	34	85.0%
High school class size	6	12.8%	5	10.0%	5	12.5%
High school rank	15	31.9%	13	26.0%	13	32.5%
High school GPA	26	55.3%	21	42.0%	21	52.5%
High school graduation date	39	83.0%	38	76.0%	38	95.0%
Prior college attended	37	78.7%	35	70.0%	35	87.5%
Transfer credit	29	61.7%	26	52.0%	26	65.0%
Remedial status/flag	15	29.8%	10	20.0%	10	25.0%
Placement test scores	16	34.0%	11	22.0%	11	27.5%
Enrollment status						
Degree-seeking status	44	93.6%	37	74.0%	26	92.5%
First term of academic history	29	61.7%	22	44.0%	22	55.0%
Full-time/part-time	47	100.0%	40	80.0%	40	100.0%
Program/major	47	100.0%	40	80.0%	40	100.0%
High school concurrent flag	24	51.1%	20	40.0%	20	50.0%
Joint enrollment flag	14	29.8%	11	22.0%	11	27.5%
Distance ed/technology flag	24	51.1%	20	49.0%	20	50.0%
Financial aid flag	30	63.8%	25	50.0%	25	62.5%
Financial aid detail	25	53.2%	20	40.0%	20	50.0%
Academic activity						
Term data collected	45	95.7%	38	76.0%	38	95.0%
Term GPA	25	53.2%	19	38.0%	19	47.5%
Term SCH attempted	45	95.7%	38	76.0%	38	95.0%
Term SCH earned	28	59.6%	22	44.0%	22	55.0%
Academic attainment						
Cumulative GPA	39	83.0%	32	64.0%	32	80.0%
Cumulative SCH earned	40	85.1%	33	66.0%	33	82.5%
Degree awarded	47	100.0%	40	80.0%	40	100.0%
Total	47	100.0%	40	80.0%	40	100.0%

Table 3: SUR database coverage of important combinations of key data elements

Data elements	SUR databases		All states		States with SUR databases	
	N	%	N	%	N	%
Core	47	100.0%	40	80.0%	40	100.0%
Set 1	44	93.6%	37	74.0%	37	92.5%
Set 2	38	80.9%	31	62.0%	31	77.5%
Set 2a	36	77.0%	29	58.0%	29	72.5%
Set 3	31	66.0%	24	48.0%	24	60.0%
Set 3a	30	63.8%	23	46.0%	23	57.5%
Set 4	22	46.8%	17	34.0%	17	42.5%
Set 4a	20	42.6%	14	28.0%	14	35.0%

Core = Sex, Race/ethnicity, Date of birth, Program/major, Degree awarded

Set 1 = Sex, Race/ethnicity, Date of birth, Geographic origin, Program/major, Full-time/part-time status, Degree awarded

Set 2 = Sex, Race/ethnicity, Date of birth, Geographic origin, Program/major, Full-time/part-time status, Credits attempted, Credits earned, Degree awarded

Set 2a = Sex, Race/ethnicity, Date of birth, Geographic origin, Program/major, Full-time/part-time status, Credits attempted, Credits earned, Cumulative GPA, Degree awarded

Set 3 = Sex, Race/ethnicity, Date of birth, Geographic origin, High school attended, High school graduation date, Program/major, Full-time/part-time status, Credits attempted, Credits earned, Degree awarded

Set 3a = Sex, Race/ethnicity, Date of birth, Geographic origin, High school attended, High school graduation date, Program/major, Full-time/part-time status, Credits attempted, Credits earned, Cumulative GPA, Degree awarded

Set 4 = Sex, Race/ethnicity, Date of birth, Geographic origin, High school attended, High school graduation date, Program/major, Financial aid, Full-time/part-time status, Credits attempted, Credits earned, Degree awarded

Set 4a = Sex, Race/ethnicity, Date of birth, Geographic origin, High school attended, High school graduation date, Program/major, Financial aid, Full-time/part-time status, Credits attempted, Credits earned, Cumulative GPA, Degree awarded

Table 4: Coverage of important combinations of key data elements by SUR database

State name	Agency responsible	Set: Core	1	2	2a	3	3a	4	4a
Alabama	SHEEO agency	X	X	X	X	X	X	X	X
Alaska	State university system	X	X	X	X				
Arizona	SHEEO agency	X	X	X	X	X	X	X	X
Arkansas	SHEEO agency	X	X						
California	Community college system	X	X	X	X	X	X	X	X
	University of California system	X	X	X	X	X	X	X	X
	State university system	X	X	X	X	X	X		
Colorado	SHEEO agency	X	X	X	X	X	X	X	X
Connecticut	SHEEO agency	X	X	X	X	X	X		
Delaware	[None]								
Florida	Statewide K-20 authority	X	X	X	X	X	X	X	X
Georgia	SHEEO agency	X	X	X	X	X	X		
Hawaii	SHEEO agency	X	X	X	X	X	X	X	X
Idaho	[None]								
Illinois	SHEEO agency	X							
Indiana	SHEEO agency	X	X						
Iowa	[None]								
Kansas	SHEEO agency	X	X	X	X	X	X	X	X
Kentucky	SHEEO agency	X	X	X	X	X	X	X	X
Louisiana	SHEEO agency	X	X						
Maine	SHEEO agency	X	X	X		X		X	
Maryland	SHEEO agency	X							
Massachusetts	SHEEO agency	X	X						
Michigan	[None]								
Minnesota	SHEEO agency	X	X	X	X				
Mississippi	SHEEO agency	X	X	X	X	X			
Missouri	SHEEO agency	X	X	X	X	X	X	X	
Montana	[None]								
Nebraska	[None]								
Nevada	SHEEO agency	X	X	X	X	X	X		
New Hampshire	[None]								
New Jersey	SHEEO agency	X	X	X	X	X	X		
New Mexico	SHEEO agency	X	X						
New York	State university system (SUNY)	X	X	X	X	X	X		
	City university system (CUNY)	X	X	X	X	X	X	X	X
North Carolina	Community college system	X	X	X	X	X	X	X	X
	State university system	X	X	X	X	X	X	X	X
North Dakota	SHEEO agency	X	X	X	X	X	X	X	X
Ohio	SHEEO agency	X	X	X	X	X	X	X	X
Oklahoma	SHEEO agency	X							

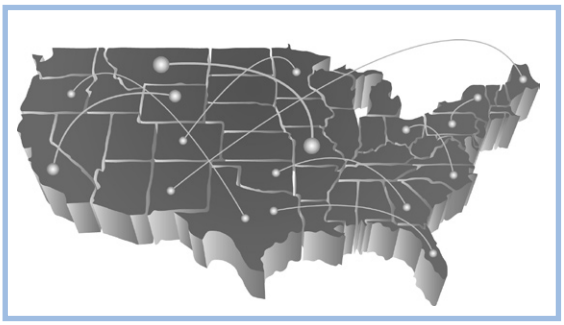
Table 4, continued

State Name	Agency Responsible	Core	1	2	2a	3	3a	4	4a
Oregon	Community college system	X	X	X	X	X	X	X	X
	State university system	X	X	X	X	X	X	X	X
Pennsylvania	[None]								
Rhode Island	[None]								
South Carolina	SHEEO agency	X	X	X	X	X	X	X	X
South Dakota	SHEEO agency	X	X	X	X	X	X	X	X
Tennessee	SHEEO agency	X		X	X	X	X	X	X
Texas	SHEEO agency	X							
Utah	SHEEO agency	X		X	X				
Vermont	[None]								
Virginia	SHEEO agency	X	X	X					
Washington	State university system	X	X	X	X				
	Community college system	X	X	X	X	X	X		
West Virginia	SHEEO agency	X	X	X	X	X	X	X	X
Wisconsin	SHEEO agency	X	X	X	X				
Wyoming	University of Wyoming	X	X	X	X	X	X	X	X
	Community college system	X	X	X	X	X	X		

Core = Sex, Race/ethnicity, Date of birth, Program/major, Degree awarded
Set 1 = Sex, Race/ethnicity, Date of birth, Geographic origin, Program/major, Full-time/part-time status, Degree awarded
Set 2 = Sex, Race/ethnicity, Date of birth, Geographic origin, Program/major, Full-time/part-time status, Credits attempted, Credits earned, Degree awarded
Set 2a = Sex, Race/ethnicity, Date of birth, Geographic origin, Program/major, Full-time/part-time status, Credits attempted, Credits earned, Cumulative GPA, Degree awarded
Set 3 = Sex, Race/ethnicity, Date of birth, Geographic origin, High school attended, High school graduation date, Program/major, Full-time/part-time status, Credits attempted, Credits earned, Degree awarded
Set 3a = Sex, Race/ethnicity, Date of birth, Geographic origin, High school attended, High school graduation date, Program/major, Full-time/part-time status, Credits attempted, Credits earned, Cumulative GPA, Degree awarded
Set 4 = Sex, Race/ethnicity, Date of birth, Geographic origin, High school attended, High school graduation date, Program/major, Financial aid, Full-time/part-time status, Credits attempted, Credits earned, Degree awarded
Set 4a = Sex, Race/ethnicity, Date of birth, Geographic origin, High school attended, High school graduation date, Program/major, Financial aid, Full-time/part-time status, Credits attempted, Credits earned, Cumulative GPA, Degree awarded

Table 5: Data element definition

Data elements	Commonality	Definition(s)
Demographics		
Sex	High	IPEDS
Race/ethnicity	High	IPEDS/census
Date of birth	High	Detail varies
Citizenship	High	Detail varies
Geographic origin	High	Detail varies
Disability status	High	Detail varies
Academic background		
Admissions test scores	Moderate	Multiple entities
High school attended	Moderate	Multiple code sets
High school class size	High	Common data set
High school rank	High	Common data set
High school GPA	Moderate	Multiple code sets
High school graduation date	High	Detail varies
Prior college attended	Moderate	Multiple code sets
Transfer credit	Moderate	Multiple code sets
Remedial status/flag	Moderate	Flag
Placement test scores	Low	Multiple entities
Enrollment status		
Degree-seeking status	High	Detail varies
First term of academic history	Moderate	Multiple code sets
Full-time/part-time	High	IPEDS
Program/major	High	Detail varies
High school concurrent flag	Moderate	Flag
Joint enrollment flag	Moderate	Flag
Distance ed/technology flag	Low	Multiple entities
Financial aid flag	High	Detail varies
Financial aid detail	Low	Multiple entities
Academic activity		
Term data collected	High	Y/N
Term GPA	Moderate	Multiple code sets
Term SCH attempted	Moderate	Multiple entities
Term SCH earned	Moderate	Multiple entities
Academic attainment		
Cumulative GPA	Moderate	Multiple code sets
Cumulative SCH earned	Moderate	Multiple entities
Degree awarded	High	Detail varies



About the authors

Peter Ewell is vice president of the National Center for Higher Education Management Systems, where he has been on the staff since 1981. His work centers on assessing institutional effectiveness and the outcomes of college. He has authored six books and numerous reports on these topics and has consulted widely with state higher education systems and individual colleges and universities. He earned a doctorate in political science from Yale University in 1976.

Marianne Boeke joined NCHEMS as a research associate in 2004. She conducts qualitative and quantitative research, does data analysis, and compiles written reports for a variety of projects. She is studying for a doctorate in higher education from the University of Denver and holds a master's degree from The American University.

Also available from Lumina Foundation for Education

Paths to Persistence: An Analysis of Research on Program Effectiveness at Community Colleges
Thomas Bailey and Mariana Alfonso
January 2005

Powerful Partnerships: Independent Colleges Share High-impact Strategies for Low-income Students' Success
Richard Ekman, Russell Garth and John F. Noonan, Editors
October 2004

Fifty Years of College Choice: Social, Political and Institutional Influences on the Decision-making Process
Jillian Kinzie, Megan Palmer, John Hayek, Don Hossler, Stacy A. Jacob and Heather Cummings
September 2004

When Saving Means Losing: Weighing the Benefits of College-savings Plans
Roberto M. Ifill and Michael S. McPherson
July 2004

Expanding College Access: The Impact of State Finance Strategies
Edward P. St. John, Choong-Geun Chung, Glenda D. Musoba, Ada B. Simmons, Ontario S. Wooden and Jesse P. Mendez
February 2004

Unintended Consequences of Tuition Discounting
Jerry Sheehan Davis
May 2003

Following the Mobile Student; Can We Develop the Capacity for a Comprehensive Database to Assess Student Progression?
Peter T. Ewell, Paula R. Schild and Karen Paulson
April 2003

Meeting the Access Challenge: Indiana's Twenty-first Century Scholars Program
Edward P. St. John, Glenda D. Musoba, Ada B. Simmons and Choong-Geun Chung
August 2002

Unequal Opportunity: Disparities in College Access Among the 50 States
Samuel M. Kipp III, Derek V. Price and Jill K. Wohlford
January 2002

Hope Works: Student Use of Education Tax Credits
Barbara A. Hoblitzell and Tiffany L. Smith
November 2001

Learning in the Fast Lane: Adult Learners' Persistence and Success in Accelerated College Programs
Raymond J. Wlodkowski, Jennifer E. Mauldin and Sandra W. Gahn
August 2001

Debts and Decisions: Student Loans and Their Relationship to Graduate School and Career Choice
Donald E. Heller
June 2001

Funding the "Infostructure:" A Guide to Financing Technology Infrastructure in Higher Education
Jane V. Wellman and Ronald A. Phipps
April 2001

Discounting Toward Disaster: Tuition Discounting, College Finances, and Enrollments of Low-Income Undergraduates
Kenneth E. Redd
December 2000

College Affordability: Overlooked Long-Term Trends and Recent 50-State Patterns
Jerry Sheehan Davis
November 2000

HBCU Graduates: Employment, Earnings and Success After College
Kenneth E. Redd
August 2000

Student Debt Levels Continue to Rise
Stafford Indebtedness: 1999 Update
Patricia M. Scherschel
June 2000

Presidential Essays: Success Stories — Strategies that Make a Difference at Thirteen Independent Colleges and Universities
Allen P. Splete, Editor
March 2000

Are College Students Satisfied? A National Analysis of Changing Expectations
Lana Low
February 2000

Fifty Years of Innovations in Undergraduate Education: Change and Stasis in the Pursuit of Quality
Gary H. Quehl, William H. Bergquist and Joseph L. Subbiondo
October 1999

Cost, Price and Public Policy: Peering into the Higher Education Black Box
William L. Stringer, Alisa F. Cunningham, with Jamie P. Merisotis, Jane V. Wellman and Colleen T. O'Brien
August 1999

Student Indebtedness: Are Borrowers Pushing the Limits?
Patricia M. Scherschel
November 1998

It's All Relative: The Role of Parents in College Financing and Enrollment
William L. Stringer, Alisa F. Cunningham, Colleen T. O'Brien and Jamie P. Merisotis
October 1998

LUMINA®
FOUNDATION FOR EDUCATION

NEW AGENDA SERIES™

Lumina Foundation for Education, a private, independent foundation, strives to help people achieve their potential by expanding access and success in education beyond high school. Through research, grants for innovative programs and communications initiatives, Lumina Foundation addresses issues surrounding financial access and educational retention and degree or certificate attainment — particularly among underserved student groups, including adult learners

Focusing on these areas, the Foundation frames issues and explores new solutions through fact-based research. Because we strive to be a credible and objective source of information on issues affecting higher education, Lumina Foundation encourages original sponsored research. The results of that research, and therefore the content of these publications, do not necessarily represent the views of the Foundation or its employees.

Believing that published research may have the longest-term impact on higher education, the Foundation publishes and disseminates articles, research reports and books. We prefer topics and approaches that are more practical than theoretical, and which emphasize pragmatic tools that will assist institutions and public policymakers.

Additional information about the Foundation's research program may be obtained from:

Dewayne Matthews
Senior Research Director
dmatthews@luminafoundation.org

Lumina Foundation for Education
New Agenda Series™

Dewayne Matthews
– Executive Editor

David S. Powell
– Director of Publications

Gloria Ackerson
Dianna Boyce
– Editorial assistance

Natasha Swingley
– Design and Production

Lumina Foundation for Education
New Agenda Series™ is published periodically by
Lumina Foundation for Education
P.O. Box 1806
Indianapolis, IN 46206-1806
www.luminafoundation.org

POSTMASTER: Send all address changes to
Lumina Foundation for Education
P.O. Box 1806
Indianapolis, IN 46206-1806.

January 2007

Helping People Achieve Their Potential®

LUMINA®
FOUNDATION FOR EDUCATION