PPIC | PUBLIC POLICY INSTITUTE OF CALIFORNIA

Increasing the Usefulness of California's Education Data

August 2013

Paul Warren and Heather Hough

Summary

Educators in California are facing a wave of new policies and practices as the 2013–14 school year begins. Schools are implementing the Common Core State Standards. The Local Control Funding Formula is expanding local fiscal flexibility and adding new K–12 resources to support services for disadvantaged students. And the community colleges are moving forward in administering the recommendations of their Student Success Task Force.

The success of these reforms will depend on the ability of local educators and administrators to translate the new policies and resources into improved instruction and student services. But the state still has important roles in the process, such as holding districts accountable and creating a system of technical assistance that builds local capacity for promoting student success.

Many states are developing state-level data systems to support the local improvement process. In a number of states, the data systems include a student-level record of achievement spanning preschool, K–12, college or university, and the workforce. For most districts and campuses, developing any type of longitudinal data system is too complex and expensive. Building a statewide longitudinal system is a more efficient approach because states already collect much of the useful data.

States with longitudinal data systems use the information to bolster the technical assistance process. Linking K–12 and higher education records also helps educators monitor and improve the transition of students from high school to college. And states can use the data to provide profiles of prospective students and diagnostic reports on the progress of students to educators and administrators. These rich databases also attract the skills of researchers, who contribute a different perspective on what works to the education dialogue.

In California, the state has made considerable progress in building its education data system. Unfortunately, it remains unfinished, with no plans for completion. Thus, a large amount of data has been collected, but K–12 data are not linked with higher education or workforce information, and much of the data remains inaccessible to educators or others who might use it to improve the functioning of the state's education programs. This report recommends a number of actions that would enable the state to quickly increase the usefulness of its data.

- **Develop ways to link preschool and K–12 data to higher education and workforce information.** Most important, linking student-level high school data with college or university records would give educators critical feedback on the success of students in their transition from the K–12 system to higher education.
- **Expand data access to educators, administrators, policymakers, and researchers.** Using its state-level data, the state should provide diagnostic reports to local educators. In addition, the state should develop various ways to access the data that meet the needs of different types of users. We also identify a change in state law that is needed to expand access to K–12 data.
- **Assign an organization or agency to coordinate access to the data and help educators use the data to improve local practices.** It makes no sense for interested users to seek out data separately from each of the five state agencies that collect it. The state needs one agency to coordinate access. This agency could also create training materials and provide technical assistance on using the information at the local level.

This approach seeks to develop California's educational data system in small steps—to minimize costs and in recognition of the fact that it will take time to determine how best to engage local teachers and administrators in using the data to improve local quality. If successful, however, this strategy will help the state maximize the value of its data in improving student outcomes.

Contents

Table

Acronyms

API	Academic Performance Index
CALPADS	California Longitudinal Pupil Achievement Data System
Cal-PASS	California Partnership for Achieving Student Success
CCC	California Community Colleges
CCCCO	California Community Colleges Chancellor's Office
CCSR	Consortium on Chicago School Research
CCSS	Common Core State Standards
CDE	California Department of Education
CORE	California Office to Reform Education
CSU	California State University
EDD	Employment Development Department
FERPA	Family Educational Rights and Privacy Act
IPEDS	Integrated Postsecondary Education Data System
LCFF	Local Control Funding Formula
PUMS	Public-Use Microdata Samples
STAR	Standardized Testing and Reporting
UC	University of California

Introduction

California is currently initiating major changes in its K–12 funding and student assessment policies. Governor Jerry Brown's Local Control Funding Formula (LCFF) in the 2013–14 state budget expands local control over the use of school funding and increases the amount of funding provided to districts with significant proportions of low-income, English Learner, or foster care students. Additionally, in 2014–15, the state will replace its current Standardized Testing and Reporting (STAR) program in grades 3–8 and grade 11 with a new assessment system in English and mathematics associated with the Common Core State Standards (CCSS).[1] At the same time, recently enacted state legislation will broaden California's system of school accountability, incorporating such measures as student readiness for college and/or employment (Chapter 578, Steinberg, Statutes of 2012).[2] The intent of these changes is to ensure that instruction, assessments, and accountability systems are aligned to support the goal of developing students who are ready for the challenges of college or the workforce.

For the most part, the state has assumed that educators have the information and knowledge to determine how best to address student needs and meet the new guidelines. Other states, however, are using data as one avenue for building local capacity to serve students well. Many states have, or are building, longitudinal educational data systems that increase the amount of information to state policymakers and local educators (Data Quality Campaign, 2012).[3] These longitudinal systems allow educators to follow students over time, compare different programs and approaches, and observe student success—not just on standardized tests, but also in college readiness and workforce participation. Several of these state systems follow students from preschool to K–12 and higher education and even through their first years in the workforce. This coordinated data system, sometimes called a P-20 database, allows educators to monitor the progress of students as they move through the educational system and enables policymakers to examine larger achievement trends, thus affording an opportunity to more effectively evaluate program success and disseminate best practices.

California has already developed high-quality data systems in K–12, and the state's public higher education institutions also have extensive data systems. But progress in combining these systems into an integrated P-20 data network has stalled. The result is that we have a large collection of data with a lot of wasted potential. K–12 data cannot be easily linked to college, university, or workforce information, and much of the data are inaccessible to educators or policymakers.

This report describes the benefits California could realize with a linked data system and the steps the state could take to make the data available and useful to educators and policymakers: we discuss the importance of a P-20 data system and how other states have capitalized on the development of such systems, examine the status of California's efforts to create such a system, and conclude with recommendations on how the state can further the development of its P-20 data system and use the data more effectively.

---

[1] California has joined the Smarter Balanced Assessment Consortium, a group of states developing new tests based on the CCSS.

[2] This legislation requires the state Board of Education to revise the Academic Performance Index (API) in such a way that test scores in high school constitute no more than 60 percent of the value of the index. The legislation does not specify the types of non-test data to be added to the API, but it does authorize the use of several performance measures, including student success in earning sufficient credits to advance to the next grade and readiness for college or employment after graduation.

[3] This national organization reported that, as of 2012, 45 states have data repositories and 14 states have databases that link K–12 education to other education programs (preschool and/or higher education).

Why a P-20 Data System Is Important

Data play a critical role in the educational improvement process. The results of state tests and real-time, locally developed student assessments help inform teachers and administrators of the progress of students and the effectiveness of individual teachers and instructional approaches. Similarly, individual schools and districts collect data to understand how students, in general, progress through their programs. Longitudinal student-level data, however, is less frequently used to inform instructional improvement: Administrative data systems are generally not designed to connect student-level information from year to year, and it is also very costly for individual districts or campuses to follow students as they move between different education systems.

This significant local cost makes the potential benefits of a statewide longitudinal database quite attractive. Data from a longitudinal data system can be shared with local educators, providing them with diagnostic information about students before they enter a program and after they leave, thus enabling faculty to tailor instructional approaches to students' needs and assess program effectiveness. Additionally, policymakers can benefit from research that uses longitudinal data to evaluate various approaches in a statewide context and can examine trends in the educational achievement of all California students. Finally, a robust P-20 data system can attract researchers who wish to conduct analyses that support the work of educators and policymakers and contributes to a larger understanding of which programs work and under what conditions. We discuss each of these applications in the following sections.

Educator Use of Data to Inform Instructional Approaches

Educators can benefit greatly from having comprehensive student information that is linked across time and across programs. In addition to the data on students that educators collect firsthand, a longitudinal data system can provide educators with another way to observe a very complex system. For example, students change: every year, some students move, graduate, or drop out, while others are identified for and receive certain programs or services. Teachers change, too: they are hired, retire, change what they teach, or move to a different grade. New instructional materials are purchased and the curriculum is refined to better reflect changing standards. In addition, time represents a critical variable in understanding the success or failure of schools. Student outcomes, for example, may depend on the quality of services delivered many years earlier.[4] A longitudinal student data system can help educators deal with this complexity. Used appropriately, such a data system can help answer questions of achievement over the long term and can also help educators analyze their programs' effectiveness in improving student outcomes.

Longitudinal data as a diagnostic tool to support student learning. Teachers and administrators can use the student information in a longitudinal database to better understand the specific problems students encounter and create new approaches in curriculum or instructional practices that lead to improved performance.

A state-level database is ideal because the independent development of longitudinal data is likely to prove to be too expensive for an individual school or program. Such an endeavor is expensive because different types of information are needed to assemble a complete picture of a student or a student body (for example, test scores, course grades, attendance, and discipline data), and several administrative offices or staff are

[4] For instance, students who are not reading at a proficient level in third grade are much more likely to drop out before graduating from high school (Hernandez, 2011).

typically involved in collecting and storing the data (Ash, 2013). This fragmentation of information at the program level complicates the effort to assemble a comprehensive record of student background and achievement. In addition, smaller schools or programs may not even have sufficient resources to coordinate these data, leaving educators on their own to figure out which data are available and how to configure the information in a useful way (Ash, 2013).

Moreover, even if a school or program is able to develop a comprehensive data system on students within its jurisdiction, it will not be able to capture information on students before they enter the school or program. This is particularly problematic in the context of student mobility in the K–12 system. It is estimated that 15 to 20 percent of school-aged children move to another school each year (Schachter, 2004); and without a statewide longitudinal data system, the schools or districts that accept mobile students often have very limited information on their prior achievement.

This is also a problem in community colleges, where campuses generally have little data on student K–12 experiences. Because prior achievement and experiences can be a strong predictor of future success, students' records from previous institutions could help teachers and faculty ensure that students are placed appropriately and have the right instructional supports to help them succeed. A P-20 data system can provide such information, which helps educators and program administrators tailor approaches to the particular students they serve.

To facilitate the effective use of longitudinal data to inform instructional approaches, many states use longitudinal data to create diagnostic reports that are sent to local educators. According to the Data Quality Campaign, 34 states use their longitudinal data to prepare reports for teachers on the growth of student achievement, while 30 states provide diagnostic reports and 20 develop "early warning" reports identifying students who may be at risk of dropping out (Data Quality Campaign, 2013).

Some states go even further. For example, Texas is piloting data "dashboards"—graphic displays of critical and timely data—that specifically respond to the needs of teachers and administrators (Ash, 2013). Georgia encourages teachers to log into the state's longitudinal data system to learn more about the prior performance of every student they teach so that they can tailor instruction to each student's needs (Georgia Department of Education, 2013).

The ideal P-20 system should provide educators with an array of data on each student's past grades, test scores, attendance, and discipline problems and also include online training modules to help educators use the data to better serve students.

Longitudinal data as a diagnostic tool to support local program evaluation. Longitudinal data can help educators track student performance while they are enrolled in a particular school, thus allowing local K–12 school districts, for example, to study how students progress through grades or to evaluate the effectiveness of new policies or programs. Individual institutions, however, rarely have access to data on students who leave, either because they graduate or they move on to another school or district. As a result, most institutions lack the ability to evaluate the effect of their programs on students' subsequent success. For example, an elementary school district would not have access to data on how its students fared in high school. Similarly, student success in post-secondary programs or the workforce are key quality indicators for school systems, but district data systems generally are not linked to higher education data or wage files. A P-20 system, however, provides the data needed to evaluate a program's effectiveness in helping students succeed after they leave the program, which in turn can improve program design.

In addition, educators often compare, with varying degrees of reliability, their student outcomes to the outcomes of students in other programs in an attempt to measure the success of their institutional performance. A comparative analysis such as this, however, depends on the availability of information on other schools that is not readily available to educators (Ash, 2013). With access to a P-20 database that spans programs and institutions, educators are able to use reliable data to evaluate whether the changes to instruction or curriculum effectively address performance problems and whether their practices are successful when measured against those of similar institutions.

State-Level Policy Research and Evaluation

State policymakers also use longitudinal data systems to answer questions about student participation in the state's programs and to ascertain the efficacy of various programs. For example:

Measuring the effects of preschool. In states with a P-20 data system, researchers are able to answer the question of how many of the state's children attend preschool and whether attendance improves achievement. In Texas, for example, researchers have used the data system to evaluate the effect of preschool education on students' elementary school success (Andrews, Jargowsky, and Kuhne, 2012). The study found that the preschool programs typically found in Texas confer long-lasting educational benefits to children.

Evaluating educational policies. States are also using their longitudinal data systems to evaluate the effects of existing or new state policies. Florida researchers used the state's comprehensive data system to show that grade retention increases student misbehavior in subsequent grades (Ozek, 2013). In another study using longitudinal data from Florida and North Carolina, researchers found that improvements in low-performing schools were driven by improvements in the productivity of the teachers and principals already in the school (Hansen, 2013).

Tracking relationships between high school completion and postsecondary success. States with data systems that span K–12 and college or university enrollment can measure the number of students who graduate from high school and college and identify the kinds of programs that improve educational attainment. For example, using longitudinal student data in Chicago and Florida, researchers found that students attending charter high schools were significantly more likely than students at non-charter public schools to graduate from high school and to attend college (Booker et al., 2011). In a study using Texas data, researchers examined the paths that students take toward a bachelor's degree to assess which paths have a higher likelihood of leading to college completion and higher future earnings (Andrews, Li, and Lovenheim, 2012). And in a study using Tennessee data, researchers showed that when students lose financial aid, it increases the likelihood of leaving college without a degree (Carruthers and Ozek, 2013).

Studying the effects of teacher quality and teacher training. Research consistently shows that teacher quality in the K–12 system is a powerful determinant of student achievement.[5] Thus, many states use longitudinal data to evaluate the effectiveness of programs designed to improve teacher quality. For example, in Louisiana, researchers used student-level K–12 data to evaluate the strengths and weaknesses of the various teacher credential programs in the state (Burns, 2010). In another study, researchers using Florida data found that advanced certification from the National Board for Professional Teaching Standards can be used to identify highly effective teachers, but that the certification itself does not enhance teacher quality (Harris and Sass, 2009). In yet another study, researchers using data from North Carolina and Florida found

[5] See, for example, Chetty et al. (2010) and Rivkin, Hanushek, and Kain (2005).

that highly effective teachers remain highly effective even if they move between schools with substantially different poverty levels or academic performance (Zeyu, Ozek, and Corritore, 2012).

As states' P-20 data systems mature, the research questions that can be answered will increase substantially; the studies highlighted above are just a small sample of the kinds of research that states can conduct to both inform state policy approaches and spread best practices.

Research Support

While P-20 data systems offer great potential for improving our understanding of the needs of students and for measuring the effects of various interventions seeking to improve the education system, these datasets are also a very complex collection of information, and they require a certain amount of skill on the part of educators and policymakers if they are to be used effectively for research and evaluation. More important, even if educators and policymakers are able to analyze the data, they may not know how to translate the data into a plan for improving curriculum and instruction. These are critical issues that need to be considered as states compile their longitudinal datasets (Data Quality Campaign, 2011).

One way to help educators and policymakers both analyze and act upon data is to develop working relationships between educators and/or policymakers and researchers. For example, a longstanding partnership between researchers at the University of Chicago and the Chicago public schools—the Consortium on Chicago School Research (CCSR)—provides actionable research findings that educators and policymakers can use to improve schools (Roderick, Easton, and Sebring, 2009). The partnership has published research on a range of topics, including the importance of school culture in student achievement and using student work to evaluate the success of schools and teachers. CCSR has also developed early warning indicators that educators are using to identify students at risk of dropping out (Allensworth and Easton, 2007).

More recently, the partnership explored reasons for the failure of many low-income students to attend college. This research highlighted the important role that high school educators can play in encouraging their students to obtain a college education (Roderick, Coca, and Nagaoka, 2011). The CCSR is an excellent example of a long-term partnership in which researchers first analyze data to shed light on services students need to succeed, and then work with educators to develop the necessary policies and programs to meet these needs.

In this same way, California can develop programs or partnerships through which researchers can provide new and useful data to educators and then work with them in analyzing the data and devising practical steps to improve student outcomes. By developing and optimizing the use of a P-20 data system, California can maximize the usefulness of its educational data.

The Status of California's Longitudinal Education Data System

Recognizing the value of a longitudinal student database, California state policymakers have made considerable progress over the past decade in developing an education data system. However, the process has stalled, the data system is only partially complete, and California is now lagging behind a number of other states engaged in similar efforts. In this section, we discuss the current status of California's education data system.

At its most basic level, a P-20 system should involve linking preschool, K–12, community college and university, and, finally, workforce outcomes. The five primary entities involved in collecting such data are K–12 school districts, the California Community College (CCC) system, the California State University (CSU), the University of California (UC), and the state Employment Development Department (EDD), which collects employment and income information on California state workers.

The development of California's P-20 data system dates back over a decade. In 2002, the state authorized the development of the California Longitudinal Pupil Achievement Data System (CALPADS), which is the cornerstone of the state's K–12 education data system (Chapter 1002, Statutes of 2002, Alpert). CALPADS collects and stores individual student data on school attendance, special program participation (e.g., English Learner and special education programs), course enrollments, and state test scores (California Department of Education, 2013). In 2008, legislation passed authorizing a more comprehensive, P-20 database (Chapter 561, Statutes of 2008, Simitian). This legislation assumed that CALPADS would be responsible for collecting data on preschool and K–12 students, and that the state's three higher education institutions (CCC, CSU, and UC) would collect student-level data from their respective campuses for use in the database.

In 2010, the state received a $6 million federal grant to begin connecting K–12 data to information on teachers collected by the Commission on Teacher Credentialing (California Department of Education, 2010). In 2011, however, Governor Brown vetoed this funding, which halted that effort (Lambert, 2011). In 2012, the governor also vetoed Senate Bill 885 (Simitian), which authorized the education entities to establish a formal agreement to implement a comprehensive student-level education data system.

The concerns motivating the governor's vetoes appear to be related to costs. His veto of funding for the teacher database in the annual budget act describes it as "a costly technology program that is not critical" (Chapter 33, Statutes of 2011). Similarly, the governor's veto message accompanying SB 885 cites "fiscal constraints."[6] However, despite these concerns, the governor permitted the development of CALPADS. Thus, while there may be other influences that contribute to the governor's skepticism about a P-20 database, it would seem that the cost of developing and operating the system is a central concern.

Despite the legislative challenges, the five state entities mentioned above are still interested in creating a statewide P-20 data system. Each of the five has developed, and maintains, databases containing administrative information on students (or, in the case of EDD, employees), and in 2011 they developed a state workgroup to collaborate on sharing data for policy and administrative purposes (California Department of Education, 2011a). However, despite the interest and commitment of these major players, the system remains incomplete.

---

[6] The governor's veto message accompanying SB 885 reads as follows: "I am returning Senate Bill 885 without my signature. This bill is unnecessary because the majority of the entities impacted by this measure have already established an interagency agreement. Should these entities choose to form a joint powers agreement in the future, they do not need additional statutory authority to do so. Whether they should or not given the current fiscal constraints—I have my doubts. Sincerely, Edmund G. Brown Jr."

As we show in the textbox below, there are four major design elements in a database that spans preschool, K–12 and college or university education, and the workforce. While the state has made considerable progress in defining the data elements and developing each entity's data system, the systems have not been linked and the data has not been made available for use. We discuss, below, the status of each step in detail.

Major Design Elements of a P-20 Education Database

COMPLETE

- **Defining data elements**: The state workgroup identified the specific data that would be merged into the P-20 database.
- **Developing each entity's data system**: With the completion of CALPADS, the data systems of the four educational agencies and EDD are capable of supplying the necessary data.

INCOMPLETE

- **Linking existing databases**: State law calls for using the K–12 student identifier to link the education data systems. There is no plan for how UC and CSU will obtain that identifier systemwide.
- **Providing access to data**: The workgroup has not developed policies or programs to share data with educators, researchers, or the public.

Defining Data Elements

The workgroup has identified the specific information to be collected in the database. The plan includes measures of academic success (courses, test scores, grades), financial aid status, and program participation (California Department of Education, 2011b). As shown in Table 1, the various entities collect a considerable amount of information on students, including demographics, program participation, and achievement.

TABLE 1
Student-level education data currently collected

Program	Data elements
Preschool[a]	Demographics[b], program eligibility[c]
Public K–12	Demographics, program eligibility, course enrollments[d], disciplinary information[e], school completion[f], state test results
Community colleges	Demographics, placement data[g], course enrollments and grades, transfer and degree completion
Public universities	Demographics, placement data, transfer credits, course enrollments and grades, financial aid data, transfer and degree completion
Employment	Quarterly wage data

SOURCES: California Department of Education, 2013b; California Department of Education, 2011c.

NOTES: a School-district-operated preschool programs only. Non–special education students are optional.
b Generally includes student's age, grade, sex, ethnicity, and school/district or campus of attendance.
c Denotes student is eligible for the federally subsidized meal program or qualifies as an English Learner, special education, or foster care student.
d High school only.
e Includes data on student offenses and disciplinary actions.
f Includes reasons students leave school and graduation information.
g Assessment information used to determine student's initial course placement in mathematics and English.

Developing Each Entity's Data Systems

There has been a substantial investment over the years to ensure that each entity's data system can supply high-quality data for the P-20 database. However, each system is in a different state of development. In 2012, after nearly a decade of work, the California Department of Education (CDE) completed the K–12 student database CALPADS, which at present includes all of California's students for the past four years. The state preschool programs and K–12 districts report these data to CDE, which compiles them in a statewide data system following students across time. However, the higher education institutions have more diffuse data collection and storage systems. The CCC Chancellor's Office (CCCCO) centrally maintains student-level databases that can provide enrollment and course data, but CSU and UC do not.[7] EDD data systems are fully equipped to track wage data over time.

Linking Existing Databases

In order to create an integrated database, the P-20 system needs a way to link the records of students who enroll in preschool, move through the K–12 system, graduate from high school, attend college or university, and subsequently join the workforce. To connect records from various points in the system, the institutions need a common student identifier. In K–12 schools, CALPADS assigns students a unique identifier. In preschools that operate through school districts, CALPADS assigns a unique identifier only to a subset of students.[8] However, the higher education databases and EDD's wage data system use social security numbers as the identifier.

There are various roadblocks to linking the data systems from different entities. For example, the link between preschool and K–12 may be difficult to establish because children attend a variety of pre-K programs, and most are not operated by school districts. At the other end of the spectrum, there is no plan for how the system would link workforce data to K–12 students who do not attend one of the three higher education institutions (CCC, CSU, or UC), although students who pursue higher education can be linked to workforce data through their social security number.

Connecting data between the K–12 system and higher education rates as the highest priority at this time, although it is not without challenges. Anticipating the need to link K–12 and higher education data, state law requires CCC, CSU, and UC to collect the CALPADS identifier for incoming students (Education Code, section 10802). However, this requirement has not worked as envisioned, since many school districts do not include the identifier on the student transcript. As a result, CSU requests students to supply their CALPADS identifier on the application for admittance. Most students, however, do not know their identifying number and leave that line on their application blank (Hirano-Nakanishi, 2013). Another approach called for the state to require districts to include the information on each student's transcript. However, that idea raised concerns about the potential costs associated with mandating the content of transcripts (Perry, 2013).

Currently, the higher education institutions are pursuing different options for obtaining the CALPADS identifier. It would be relatively easy for CSU and UC to collect the identifier on a student transcript, since

---

[7] Individual CSU and UC campuses do collect this information, but the extent to which it is maintained in a central location varies. CSU and UC centrally maintain student-level demographic and program participation data, but not individual student enrollment records (Perry, 2013).

[8] Per CDE, all local education agencies operating child development programs or preschools must assign a student identifier to special education students and are encouraged (but not required) to do so for all other students. For more information, see www.cde.ca.gov/ds/sp/cl/systemdocs.asp, CALPADS Statewide Student Identifier (SSID) and Enrollment Procedures (DOC; Dated 01-Mar-2012).

these institutions require students to include a transcript with their applications. Given the extensive amount of information on student coursework and grades that is currently reported on transcripts, it would seem to be a relatively trivial matter to include a student identifier. Districts are not obligated to provide the identifier, however, and the lack of universal compliance could compromise the data system. Given this situation, CSU initiated a pilot program involving an electronic transcript that provides the CALPADS identifier as well as course work and grades (Forbes, 2013).

The CCC Chancellor's Office is heading in a different direction. Since the CCCs maintain a policy of open enrollment, they do not obtain high school transcripts. So the CCC is seeking to obtain the identifiers by submitting student names and demographic information to the CALPADS locator service. This is the same process schools and districts use to obtain the CALPADS records of students who move from one district to another within California. The chancellor's office is currently working with the administrator of the locator service to implement the matching process in order to determine whether it generates a sufficient matching rate for the CCCs (Perry, 2013).

Providing Access to Data

The fourth component of a P-20 database is access—creating pathways that enable different users to obtain and use the longitudinal information. The Data Quality Campaign encourages states to develop a variety of policies and data products that match the interests and sophistication of different types of users. The group concludes that California has taken few actions to make data accessible and useful (Data Quality Campaign, 2012). For example, if K–12 teachers and administrators are considered to be some of the main users of the data system, information must focus on the issues they care about and be shared in ways they can apply to their schools. Similarly, to meet the needs of researchers and policymakers, the state needs to balance the release of student-level data with the need to protect individual student information from exposure. For the most part, these issues have gone unattended, and the data largely represent an untapped resource.

With so many education policies changing in California—the introduction of the Common Core State Standards, changes to our existing systems of school accountability, the Local Control Funding Formula—researchers and policymakers will need to assess the effects not only on students but on a range of other social and political stakeholders as well, including teachers, schools, local residents, and taxpayers. While California currently realizes only a small return on the investments it makes in collecting data, this is easily remedied. In the next section, we discuss the steps the state could take to further develop its educational data system and help researchers, educators, and policymakers at all levels use the data to improve the quality of education in the state.

Recommendations

Although California has undertaken a considerable amount of work in developing the data collection capacity among different entities spanning preschool, K–12, college, and the workplace and, in the process, amassed a large collection of data, it has yet to develop a connected and coordinated system of data. The question is, how can the state now proceed to make the data more useful?

Our recommendations fall into three general areas:

- The five state agencies should develop a way to connect preschool and K–12 data to higher education and workforce information.
- The state should increase access to its data.
- The education entities and EDD should identify an organization or agency responsible for coordinating access to the data and helping educators use the data to improve local practices.

The goal of our recommendations is to help the state move toward a low-cost plan for quickly increasing the value of the state's education data. As a result, this plan falls short of a comprehensive P-20 database in that it would leave the existing system largely intact, with each education system responsible for collecting and storing its own data. However, by ensuring reliable links between K–12, higher education programs, and the workplace, educators and policymakers would be able to obtain better information on the transition of students through various educational institutions and even into employment after graduation. In addition, the state could begin to reap a return on the investment it has made in collecting student-level data by promoting the use of the data to improve local programs. Furthermore, creating reliable links among various P-20 entities will lay the groundwork for an integrated data system in the future. We discuss our recommendations in greater detail below.

Recommendations for Increasing the Usefulness of Existing Education Data

Develop a longitudinal data record from preschool to the workforce.

- Require districts to include the CALPADS identifier on student transcripts.
- Develop plans for expanding the collection of preschool data and for linking higher education data with workforce outcomes (EDD records).

Increase data access for educators, policymakers, and researchers.

- Promote the use of longitudinal data as a resource for educators and administrators.
- Allow support organizations to access K-12 data directly from CALPADS, thereby reducing costs and increasing the effectiveness of the program.
- Revise state privacy policies to be consistent with federal law.

Create an administrative group responsible for managing the data.

- Establish policies and data products that allow access to data by a wide range of users while also protecting the privacy of students.
- Develop data tools that help local educators make use of longitudinal data.
- Create training materials for translating data and research into improved practice.

Develop a Longitudinal Data Record from Preschool to the Workforce

The most important developmental step for the system would be to establish a way to connect student records, beginning in preschool and concluding with EDD employment and wage data, thereby creating a longitudinal data record that would follow students as they progress through the system. Linking data from K–12 to the state's colleges and universities would provide immediate benefits.

As discussed above, the existing data system cannot provide college-success data to districts that want to evaluate the effectiveness of their programs in helping students after they graduate. Currently, a K–12 school district interested in studying how its students fare in college is faced with the challenge of linking its students to higher education systems. To do that, each district would need to develop methods for linking students across time and programs. In addition, each district would have to establish an agreement with the local CSU or UC (or with EDD for workforce data). If linked data were available, these steps would already have been completed, streamlining the process.

An integrated system would also benefit California's higher education institutions. The community colleges and CSU, for instance, would like access to K–12 data to strengthen the placement process and reduce remediation rates of entering freshmen (Hirano-Nakanishi, 2013). Again, such a relationship can be established, but as it stands now, each higher education institution would need to develop a data sharing agreement and procedure with each K–12 district that feeds students into the community college or university. Several examples of this kind of relationship exist—for instance, the Long Beach Community College has partnered with the Long Beach Unified School District to share student information (CSU Long Beach, 2013)—but not many campuses have developed this kind of partnership.

From an administrative point of view, improvements in the data system could reduce the burden of federal and state reporting requirements that involve multiple education entities (Perry, 2013)—for example, linking graduates to EDD wage files for federal vocational education accountability reports and tracking community college rates of transfer to CSU and UC.

One of the key issues in connecting the different pieces of the database is whether to use social security numbers or the CALPADS student identifier as the linking device. Social security numbers have several advantages. Most students have Social Security numbers, and districts currently collect and maintain them for at least some of their students.[9] Since the higher education institutions and EDD use Social Security numbers as the student identifier, using the Social Security number as an identifier for preschool and K–12 students would make linking the data relatively simple.

State law, however, reflects the legislature's reluctance to require the use of Social Security numbers in K–12 education. The Education Code directs CALPADs to develop its own identifier system. In addition, state law requires using the CALPADS identifier to link K–12 and higher education data. This makes connecting the data with higher education more difficult and makes linking K–12 data to EDD wage data virtually impossible. At this juncture, rather than reopening the debate on the choice of student identifiers, we think it makes more sense to focus on connecting K–12 and higher education data using the CALPADS identifier. As the benefits of the data system become more tangible, the issue of linking K–12 and EDD data (and the use of the Social Security number as an identifier) could be revisited.

[9] Education Code section 56601, for instance, requires districts to collect social security numbers for a special education database.

Including an identifier on student transcripts. In using the CALPADS identifier, however, the higher education agencies must solve the problem of somehow obtaining the number through a student's high school transcript. Thus, the most obvious and easiest way for UC and CSU to obtain the K–12 identifier would be to require that districts include the identifier on transcripts. This could be done in several ways. The state could simply require that transcripts include the CALPADS identifier and, if need be, reimburse districts for the small cost of this activity.

Alternatively, the state could tie the identifier issue to the larger need to convert transcripts from a paper to electronic format. Currently, most transcripts are provided on paper, and UC and CSU convert them to a virtual format. The state could require districts to send transcripts (including the identifier) to UC and CSU electronically. Since a virtual format saves the universities the cost of transcription, this option would save money at the higher education level. To make this option workable for all 1,000 K–12 districts (and to minimize the costs to districts), UC and CSU would need to develop formats that could be easily implemented at the local level.

Developing plans for linking education and workforce data. Linking higher education records with workforce data from the Employment Development Department would facilitate an even more comprehensive analysis of the success of students as they complete educational programs in California. Linking with the EDD data is relatively straightforward since the higher education systems and the EDD system both use social security numbers as unique identifiers. Indeed, the community college system in California has been connecting student records to EDD records in this way for many years. Several UC and CSU campuses have recently established links with EDD data as well, and some efforts have been made to integrate this data centrally.[10]

However, the state would still need to resolve the longer-term issue of including all preschool students in the database and connecting K–12 records to EDD data for students who do not pursue postsecondary education. Regarding the preschool issue, the state could issue CALPADS identifiers for all students in district-administered preschool programs. While this would only include students in preschool programs associated with school districts (a minority of all students enrolled in preschool programs), it would be a first step and would provide experience that would help determine whether it is cost-effective to expand the types of preschool students included in the database. As discussed above, linking K–12 records to EDD data is an issue to be addressed at a later date.

Increase Data Access

Another way to capitalize on the existing data is to make them more accessible to educators and researchers. As it currently stands, each school district or program sends a great deal of data to the state, but these data are not "shared back" with the sending districts or others in any way that supports meaningful analysis. In the K–12 system, for example, each school district sends its student-level data each year to CALPADS, which CALPADS uses to create a K–12 longitudinal data system. In addition to merging these files over time, CALPADS also provides the very valuable step of following students who move between districts in the state. Since each school district provides its own data to CALPADS, the districts clearly have the data on

[10] For example of local efforts, see Cal State Northridge: www.calmis.ca.gov/file/Advisory-Group/Handouts/Jan-12/CASUN-WhatEverHappenedtoPresentation.pdf. For example of how UC integrated EDD data for all UC graduates, see www.cair.org/conferences/cair2012/pres/38_Furgiuele.pdf.

their own students. However, many districts lack the resources or expertise to construct these data into a useable, longitudinal database.

This problem is especially acute for smaller districts. Large districts usually have the in-house resources to support in-depth research and even attract the attention of external researchers. Several large urban districts participate in partnerships in which researchers support the development and analyses of longitudinal data systems. The University of California, San Diego, for example, has teamed up with the San Diego Unified School District to examine issues that contribute to decisionmaking in the San Diego district. Similarly, a new research institute has recently been established in Los Angeles to support the effective use of data and research in school improvement (Los Angeles Education Research Institute, 2013), and Stanford University and the San Francisco school district have a similar partnership.[11] Even these research partners are often limited to data that they can collect from the institutions with which they work, since links to other K–12 or higher education data are more difficult to establish.

Despite these exemplars, most school districts across the state are smaller and generally lack the capacity to develop longitudinal data systems, conduct analyses with them, or leverage them to establish connections with other institutions. Every connection to other K–12 districts or to CSU, CCC, UC, or EDD represents a layer of coordination that most systems do not have the manpower to handle. Thus, every step the state can take to reduce the local barriers to using the longitudinal data makes the data more valuable for improving our educational system.

Promote the use of longitudinal data to improve education outcomes. The state should use its database to help teachers, administrators, researchers, and policymakers improve the educational environment and student outcomes. CDE currently releases school- and district-level data through its website, but the site limits its reports to specific data and does not attempt to provide the kind of diagnostic reports that would make the data particularly useful to school or district professionals. In addition, its data portal reports only K–12 data and does not use CALPADS's multiyear perspective to report on student growth or success in higher education.[12] There are a number of ways in which the state can make data more useful to local educators.

To begin with, the state can develop ways of releasing longitudinal data that are accessible and useful to teachers and administrators. Each education institution could develop reports and data tools that help educators use the data to evaluate student progress. The CCCCO, for instance, recently published a "salary surfer" that uses EDD data to report on the earnings of recent CCC graduates by program (CCCCO, 2013). As discussed above, other states use longitudinal data to develop diagnostic analyses of student test results and timely reports for K–12 teachers on the growth of student achievement and the success of high school graduates in colleges or universities.

The state can also improve practices by making its data accessible to researchers. As noted above, states with robust longitudinal data systems are attractive to researchers because of the richness of the questions that can be answered. A true P-20 statewide database in California would attract world-class researchers to conduct analyses on program and policy effectiveness, which would greatly contribute to our state's collective knowledge of the kinds of approaches that work, as well as those that do not.

---

[11] For more information on the Stanford /SFUSD Partnership, see http://silvergiving.org/stanfordsfusd-partnership.
[12] There is one exception to this: The CDE website reports student success in post-secondary education for high school students who graduated in 2006–07. These data were developed with one-time federal funding.

One way to support the capacity for data analysis, on the part of both educators and policymakers, is through the development of research partnerships. For instance, the state could create a research partnership to support district efforts to improve programs for disadvantaged students, consistent with the recent fiscal overhaul. Such state-sponsored partnerships could help address the research needs of small school districts or colleges, which have a hard time developing partnerships with external researchers. These relationships represent a potentially low-cost way for the state to take advantage of the desire of researchers to contribute to improving student achievement. In fact, the state may be able to establish partnerships at almost no additional cost: allowing researchers access to the data may provide sufficient means to enable them to obtain financial support for the work.

Allow support organizations to obtain data through the state's longitudinal system. The state should also support other efforts that can assist educators in using data to improve the educational system. For example, the state's comprehensive longitudinal data system would be extremely useful to organizations such as the California Partnership for Achieving Student Success (Cal-PASS), which seeks to help institutions in the K–16 system use data to raise student achievement, close achievement gaps, and increase college readiness and success across the state. The data would also be useful to networks of educational institutions that are working together to improve outcomes for their students. One such group is the California Office to Reform Education (CORE), a collaboration of ten California school districts.[13] With access to an integrated data system, these organizations could leverage data analyses to help diagnose student problems and study the effectiveness of different approaches. However, as it stands now, organizations like CORE and Cal-PASS face the same data constraints as local institutions, which limits their ability to assist local districts with data analysis or provide cross-pollination across different institutions.

To increase the effectiveness of Cal-PASS, CORE, and other groups with similar goals, the state could authorize them to directly access the data system. Direct access to the data would help Cal-PASS in two ways: First, it would permit the organization to be more efficient. Currently, Cal-PASS collects and cleans the data it receives from local educational institutions—repeating the process the state goes through when it adds information to its database. This duplication of effort seems unnecessary.

Second, the state's databases would enable Cal-PASS to target its services more effectively. Currently, Cal-PASS works with interested campuses and districts to improve the alignment of K–12 and community college standards. With CALPADS data, however, Cal-PASS could become proactive—using the statewide files to identify regions in which the alignment between K–12 and higher education looks particularly problematic. Thus, access to the K–12 data would allow Cal-PASS to focus on the areas of greatest need for its services. Of course, Cal-PASS would have to follow the protections for individual student records as required by the Family Educational Rights and Privacy Act (FERPA), and it would have to develop clear procedures for handing data that satisfies both CDE and the higher education institutions.

Clarify and moderate the state's student-testing privacy policy. Certain provisions of state law create an unnecessary barrier to sharing student data. Specifically, the Education Code generally prohibits the state from releasing individual scores on the Standardized Testing and Reporting (STAR) tests (Education Code, section 60607).[14] All other individual K–12 pupil data—including student scores on other state tests as well as the higher education data managed by the state—are protected under FERPA. The federal law provides

[13] For more information on the CORE districts, see http://coredistricts.org/.
[14] The STAR program tests all students in mathematics and English in grades 2–11 each year. Science and history tests are administered in grades 9–11 and in selected earlier grades.

detailed guidelines for protecting student privacy, but it does not prohibit states from releasing student test data for legitimate administrative or research purposes (United States Department of Education, 2013).

Other areas of state law appear to encourage providing access to K–12 data. In conflict with the STAR privacy language, section 49079.6 of the Education Code authorizes CDE to release pupil data to researchers if they can satisfy the requirements of FERPA. According to CDE, however, the STAR privacy language takes precedence. As a result, CDE believes that the state law allowing release of data to researchers does not apply to STAR data (Ashley, 2013).

The state privacy language also prohibits other state and local education entities from accessing the STAR data. CSU, for example, would like to be able to access student test results to identify promising students. Community colleges could use STAR data for placement decisions. Because the STAR data constitute the core of the K–12 portion of a longitudinal data system, changing this state policy is a critical step in making the educational data more useful.

Current legislation (Assembly Bill 484, as amended July 3, 2013) would extend the state privacy rules to the results of the new tests of the Common Core State Standards. If the state wants to maximize the value of its education data, it needs to reconsider this policy. We recommend that the legislature harmonize the student-testing privacy protections with those of FERPA.

Create an Administrative Group Responsible for Managing the Database

Once data from the various entities can be linked at the student level, the state will have developed a decentralized data system in which each entity separately collects and stores data. While this represents a critical step in the creation of an integrated system, it does not resolve the issue of how interested users could access the data. Since, in this short-term plan, there would be no single database containing the information from the institutions composing the system, potential users would have to separately request data from each entity. Moreover, while the data would become more accessible, it is likely that, without training or assistance, some potential users would lack the skill to analyze the data and use the results to improve local programs.

To address these gaps, the four education entities and EDD should create an administrative group responsible for overseeing the database and arranging access to the data while also ensuring the privacy of individual students. The administrative group could also develop a protocol with the state educational entities for satisfying data requests, and begin the process of developing materials to help local educators analyze and use longitudinal data to inform practice and increase student achievement.

With regard to the structure of this group, we see several possibilities. The five state agencies involved in collecting and maintaining the data could jointly administer the data base operations and training effort or assign the duties to one entity in the group. Alternatively, these activities could be vested in a separate and distinct administrative group guided by the five entities, similar to Cal-PASS. What seems more important than the structure of the administrative group is the commitment of the educational entities to promoting the use of the data for improving the quality of the state's education system. The initial efforts of the administrative group should focus on the three priorities discussed below.

Coordinate access to K–12 and higher education data. First, as noted above, there is no one place where a researcher can gain access to K–12 and higher education data files. Those interested in monitoring the

transition from high school to college and university, for instance, must approach each institution separately, navigating the different procedures and policies currently in place. The administrative group should develop policies and timelines that simplify this process, streamline access to data, and clarify the privacy policies that must be followed.

Develop new data products to expand access. The educational information in the various databases is currently available only in the form of year-by-year student-level files at the institutional level, and the technical skills needed to analyze these data represent a barrier to their use for purposes of policy and practice. However, in many cases, accessing student-level data is unnecessary—policymakers and other analysts can use aggregated or simplified versions of longitudinal data to meet their needs. Thus, the administrative group should develop new data products that "translate" the student-level files into longitudinal information accessible to a broad range of users.

These data products could take several forms. The Census Bureau, for instance, issues a "public use" file of individual census responses, which allows users to analyze a sample of representative households.[15] A similar sample of student-level files could be developed using the state's educational data. This would make longitudinal data available to virtually any interested user while guaranteeing the privacy of individual student information.

Alternatively, the data could be accessed through a simple data tool that allows users to analyze the information without actually taking possession of the student-level files. There are several excellent examples of this kind of data access throughout the state and also nationwide. For example, the Integrated Postsecondary Education Data System (IPEDS) administered by the National Center for Education Statistics has a web portal that allows users to explore and analyze institutional and survey data. In California, the CCC Chancellor's Office runs a "data mart" that provides access to a limited collection of community college data.

Create training materials for translating data into improved practice. In addition to making the data more accessible, the administrative group should develop training materials that will help educators analyze the data and understand the ways in which they can use it to improve practices and programs that will lead to better student outcomes. The administrative group could work with CDE and the higher education entities to ensure that the capacity-building materials meet the needs of different users in schools, colleges, and universities. Over time, this role could expand if local users needed a higher level of support to take advantage of the data.

[15] The samples, known as Public-Use Microdata Samples (PUMS), are available at www.census.gov/main/www/pums.html.

Conclusion

The state has made considerable progress in building an educational data system. Unfortunately, it remains unfinished. We have a large accumulation of data, but they remain inaccessible to most of those who could use the information to improve the quality of education in California. To put it another way, the data represent a tremendous potential within the state system to help local districts and schools improve student outcomes. We should not let this potential remain unrealized.

This report identifies the needed next steps for the system. The goal of our recommendations is not to create a *finished* P-20 data system, but to move the system in a direction that would make the data more useful in improving the state's education system. Making the data more accessible and useful to educators, for example, would help generate a return on the time and expense that has already been invested in collecting the student-level data. And connecting K–12 and higher education data could help educators improve the transition from high school to college or university.

Our strategy of taking small steps is purposeful—it allows for refinements in the data system at a pace and scope that minimize costs. It also allows time for the education entities—the high schools, colleges, and universities of California—to determine how best to engage local teachers and administrators in using the education data. It also gives time for all parts of the education system to evaluate the benefits from the state's data system. If successful, this strategy will help the state maximize the value of its data in improving student achievement. This would be a big step forward, even if it means that a more comprehensive, integrated P-20 data system will be a work in progress for some time to come.

References

Allensworth, Elaine M., and John Q. Easton. 2007. *What Matters for Staying On-Track and Gradating in Chicago Public High Schools, A Close Look at Course Grades, Failures, and Attendance in the Freshman Year.* Consortium on Chicago School Research, University of Chicago.

Andrews, Rodney J., Paul Jargowsky, and Kristin Kuhne. 2012. "The Effects of Texas Targeted Pre-Kindergarten Program on Academic Performance." Working Paper 84, National Center for Analysis of Longitudinal Data in Education Research.

Andrews, Rodney J., Jing Li, and Michael F. Lovenheim. 2012. "Heterogeneous Paths Through College: Detailed Patterns and Relationships with Graduation and Earnings." CALDER working paper. Available at www.caldercenter.org/publications/upload/working-paper-83.pdf.

Ash, Katie. 2013. "Fragmented Data Systems Are a Barrier to Better Schools, Expert Says." *Education Week*, March 11, 2013.

Ashley, Keric. 2013. Director: Analysis, Measurement, and Accountability Reporting Division, California Department of Education. Personal communication.

Booker, Kevin, Tim R. Sass, Brian Gill, and Ron Zimmer. 2011. "The Effects of Charter High Schools on Educational Attainment." *Journal of Labor Economics* 29 (2): 377–415.

Burns, Jeanne M. 2010. "Using Teacher-Student Data to Evaluate the Effectiveness of Teacher Preparation Programs in Louisiana." American Youth Policy Forum and Data Quality Campaign, September 28, 2010.

California Community Colleges Chancellor's Office. 2013. Student Success Initiative, Salary Surfer, retrieved at http://salarysurfer.cccco.edu/SalarySurfer.aspx.

California Department of Education. 2013. CALPADS Background. Available at www.cde.ca.gov/ds/sp/cl/background.asp.

California Department of Education. 2013b. CALPADS Code Sets, Version # - 4.1, April 23, 2013. Available at www.cde.ca.gov/ds/sp/cl/systemdocs.asp.

California Department of Education. 2011a. Interagency Agreement Between the California Department of Education, California Community Colleges Chancellor's Office, California State University Chancellor's Office, University of California Office of the President, State of California Employment Development Department, and the California Postsecondary Education Commission.

California Department of Education. 2011b. Application for the California Comprehensive Education Data System.

California Department of Education. 2011c. Preliminary P-20 Data Elements, California Comprehensive Education Data System. In California Department of Education, Institute for Education Sciences State Longitudinal Data System Application, December 2011.

California Department of Education. 2010. California Longitudinal Pupil Achievement Data System Update, October 5, 2010. Available at www.cde.ca.gov/be/pn/im/documents/infooct10item04.doc.

California State University at Long Beach. 2013. "Long Beach Education Partnership: Seamless Education at Its Best." Available at www.csulb.edu/president/education-partnership.

Carruthers, Celeste K., and Umut Ozek. 2013. "Losing Hope: Financial Aid and the Line Between College and Work." CALDER working paper. Available at www.caldercenter.org/publications/upload/wp91.pdf.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and. D. Yagan. 2010. "How does your kindergarten classroom affect your earnings? Evidence from project STAR." National Bureau of Economic Research.

Data Quality Campaign. 2013. What does it take for states to accomplish DQC's Action 6: Create progress reports using individual student data to improve student performance. Available at www.dataqualitycampaign.org/stateanalysis/actions/6/.

Data Quality Campaign. 2012. California 2012 DQC State Analysis. Available at www.dataqualitycampaign.org/stateanalysis/states/CA/actions/.

Data Quality Campaign. 2011. Using Data to Improve Teacher Effectiveness: A Primer for State Policymakers." Available at www.dataqualitycampaign.org/files/DQC-TE-primer-July6-low-res.pdf.

Forbes, Eric. 2013. Assistant Vice Chancellor, California State University. Personal communication, May 15, 2013.

Georgia Department of Education. 2013. *Statewide Longitudinal Data System Online Teacher Training, Who Are My Students?* Available at http://cms.gavirtualschool.org/Development/SLDS/teacher/SLDS002/index.html.

Hansen, Michael. 2013. "Investigating the Role of Human Resources in School Turnaround: A Decomposition of Improving Schools in Two States." CALDER working paper. Available at www.caldercenter.org/publications/upload/wp89.pdf.

Harris, D. N., and T. R. Sass. 2009. "The effects of NBPTS-certified teachers on student achievement." *Journal of Policy Analysis and Management* 28 (1): 55–80.

Hernandez, Donald J. 2011. *Double Jeopardy: How Third-Grade Reading Skills and Poverty Influence High School Graduation.* Annie E. Casey Foundation.

Hirano-Nakanishi, Marsha. 2013. Assistant Vice Chancellor, Academic Research and Resources, California State University. Personal communication. March 12, 2013.

Lambert, Diana. 2011. "Loss of Teacher Data Grant Could Affect State Chances for Federal Education Funding." *Sacramento Bee*, August 11. Available at www.sacbee.com/2011/08/11/v-mobile/3830457/loss-of-teacher-data-grant-could.html#storylink=cpy.

Los Angeles Education Research Institute. 2013. "Welcome to LAERI." Available at www.laeri.org/.

Ozek, Umut. 2013. "Hold Back to Move Forward? Early Grade Retention and Student Misbehavior." CALDER working paper. Available at www.caldercenter.org/publications/upload/WP-100.pdf.

Perry, Patrick. 2013. Vice Chancellor of Technology, Research, and Information Systems, California Community Colleges Chancellor's Office, Personal communication, April 19, 2013; July 24, 2013.

Rivkin, S., E. Hanushek, and J. Kain. 2005. "Teachers, schools, and academic achievement." *Econometrica* 73 (2): 417–458.

Roderick, Melissa, Vanessa Coca, and Jenny Nagaoka. 2011. "Potholes on the Road to College: High School Effects in Shaping Urban Students' Participation in College Application, Four-year College Enrollment, and College." *Sociology of Education* 2011 (84): 178.

Roderick, Melissa, John Q. Easton, and Penny Bender Sebring. 2009. *The Consortium on Chicago School Research: A New Model for the Role of Research in Supporting Urban School Reform*. Consortium on Chicago School Research, University of Chicago.

Schachter, Jason. 2004. *Geographic Mobility: Population Characteristics March 2002 to March 2003*. United States Census Bureau.

United States Department of Education. 2013. Family Educational Rights and Privacy Act (FERPA). Available at www.ed.gov/policy/gen/guid/fpco/ferpa/index.html.

Zeyu Xu, Umut Ozek, and Matthew Corritore. 2012. "Portability of Teacher Effectiveness Across School Settings." CALDER working paper. Available at www.caldercenter.org/publications/upload/wp77.pdf.

About the Authors

Paul Warren is a policy associate at PPIC, where he focuses primarily on K–12 education finance and accountability. Before he joined PPIC, he worked in the California Legislative Analyst's Office for more than twenty years as a policy analyst and manager. Previously, he was chief consultant to the State Assembly's committee on education. He also served as deputy director for the California Department of Education, helping to implement current testing and accountability programs. He holds a master's degree in public policy from Harvard's Kennedy School of Government.

Heather J. Hough is a research associate with the Carnegie Foundation for the Advancement of Teaching, where she focuses on helping educators use research to drive improvement. Previously, she was a research fellow at PPIC, where she focused on teacher compensation, support, and accountability. She has worked as a researcher in the Center for Education Policy Analysis at Stanford University and the Center for Education Policy at SRI International. She holds a Ph.D. in education policy from Stanford University.

Acknowledgments

The authors would like to acknowledge the time and assistance of Patrick Perry of the California Community Colleges Chancellor's Office, Keric Ashley and Paula Mishima of the California Department of Education, Marsha Hirano-Nakinishi of the California State University Chancellor's Office, and Jim Lanich and Ken Sorey of Cal-PASS in helping us understand the issues and history of the state's longitudinal education data system. The report benefited from the comments and feedback of Laura Hill, Hans Johnson, David Plank, and Brain Sala. Lynette Ubois and Gary Bjork provided excellent editorial contributions.

Research publications reflect the views of the authors and do not necessarily reflect the views of the staff, officers, or Board of Directors of the Public Policy Institute of California. All errors are our own.

PPIC | PUBLIC POLICY INSTITUTE OF CALIFORNIA

The Public Policy Institute of California is dedicated to informing and improving public policy in California through independent, objective, nonpartisan research on major economic, social, and political issues. The institute's goal is to raise public awareness and to give elected representatives and other decisionmakers a more informed basis for developing policies and programs.

The institute's research focuses on the underlying forces shaping California's future, cutting across a wide range of public policy concerns, including economic development, education, environment and resources, governance, population, public finance, and social and health policy.

PPIC is a private operating foundation. It does not take or support positions on any ballot measures or on any local, state, or federal legislation, nor does it endorse, support, or oppose any political parties or candidates for public office. PPIC was established in 1994 with an endowment from William R. Hewlett.

PUBLIC POLICY INSTITUTE OF CALIFORNIA
500 Washington Street, Suite 600
San Francisco, California 94111
phone: 415.291.4400
fax: 415.291.4401
www.ppic.org

PPIC SACRAMENTO CENTER
Senator Office Building
1121 L Street, Suite 801
Sacramento, California 95814
phone: 916.440.1120
fax: 916.440.1121