Measuring College Performance

Lessons for Policymakers

Erica Blom, Kristin Blagg, Matthew M. Chingos, Tomas Monarrez, Macy Rainer, and Kelia Washington

January 2020

The availability of US higher education data has exploded over the past decade. Nearly every state has a student-level longitudinal data system that follows students from the K–12 system into college (and often the workforce), something only 12 states could do in 2005.[1] A national data system does not yet exist, but the federal government releases institution- and program-level data on average earnings and debt among former college students through the College Scorecard and has collaborated with academic researchers to produce college-specific measures of economic mobility.[2]

The production of new data has focused on the needs of potential students and their families, but the available metrics are often not as useful to students as policymakers hope.[3] For example, institution-level measures are of limited use to students whose college choices are constrained by family obligations that keep them close to home or academic credentials that limit their choices to less-selective institutions (Blagg and Chingos 2016).

State policymakers could also benefit from better data on college performance, but their needs may differ from students' needs. A student may benefit from knowing how students like her fare at particular colleges, whereas state legislators may want to assess the overall contribution of each public college to the state's educational attainment and economic development. And a measure that is useful to one audience may be misleading to another. A highly subsidized public college may be a great financial deal for students because of the low relative price, but state policymakers could see the college as a poor investment relative to other institutions that would make more efficient use of the same public resources.

The way state policymakers use student-level data systems tends to be limited to regular reporting of summary statistics and compliance with state and federal law. For example, the website of the

Colorado Student Unit Record Data System lists the following examples of how the data are used: providing institution-level enrollment counts and information on student characteristics, reporting Integrated Postsecondary Education Data System data to the federal government, calculating graduation rates, providing data for an annual diversity report, and completing program-level projections.[4]

Several issues must be addressed to enhance the production and use of higher education metrics. Existing metrics are calculated at different levels of aggregation (e.g., institution versus program), are sometimes adjusted for measures of student preparation (using different methodologies), and draw on a wide variety of outcomes (e.g., degree completion, earnings, and student loan repayment) that are based on different subgroups of students (e.g., full-time students, graduates, and full-time workers). And it is often difficult to separate college quality from differences in students' academic preparation and background characteristics.

Urban Institute researchers worked with state policymakers in Connecticut and Virginia to better understand their needs for higher education data and make progress on some of these thorny technical issues. In this brief, we summarize lessons learned for this body of work.

Lesson 1: Use Student-Level Data to Measure How Well Colleges Serve Students.

Graduation rates are commonly used to judge college performance, but they can confuse institutional quality with student characteristics. For example, selective colleges that only admit students with strong academic preparation are likely to have better outcomes than open-access colleges, regardless of the quality of instruction they provide. Looking only at the graduation rate is likely to give policymakers misleading impressions.

An alternative to raw performance measures is to develop "value-added" or "adjusted" measures that provide a more nuanced view of an institution's performance by taking student characteristics (e.g., demographics and academic preparation) into account. These adjusted measures can help state policymakers determine the relative efficacy of their state's higher education institutions. If adjusted measures alter the rankings relative to raw measures, that could affect conclusions about how certain institutions or programs are performing or how to allocate resources.

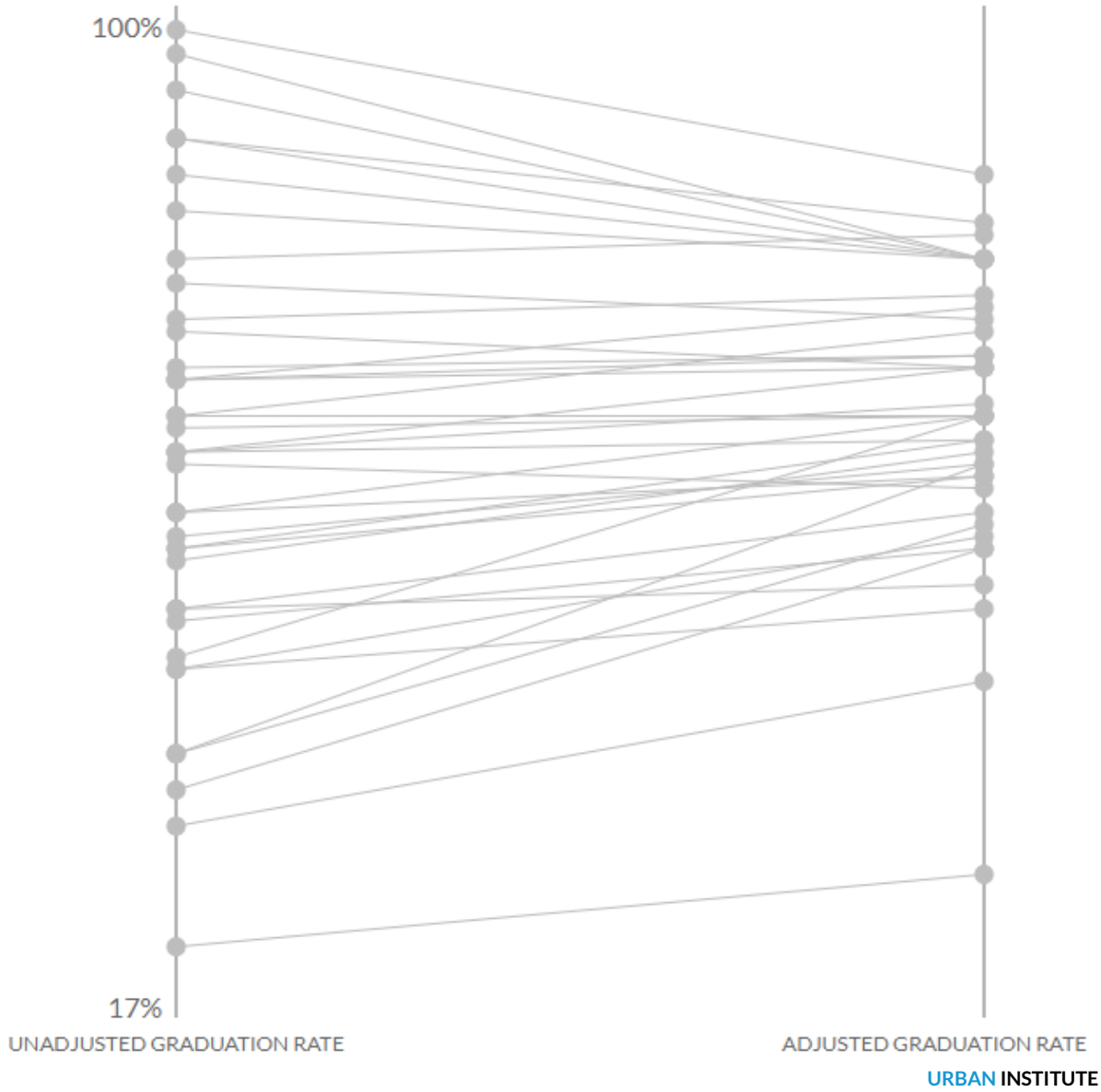
Based on data from Connecticut and Virginia, we find that student demographics and academic preparation explain a considerable amount of the differences in graduation rates. At four-year colleges, the adjustments reduce the distribution of graduation rates by as much as 20 percentage points and alter the ranking of institutions (figure 1). Differences in academic preparation, rather than demographics, account for most of these adjustments.

At two-year colleges, the overall adjustments are smaller, and race, ethnicity, and family income play a larger role than academic preparation. This likely reflects the fact that community colleges

primarily enroll students from the surrounding area, whereas four-year colleges vary in their selectivity, as they admit students from around the state.

FIGURE 1

Unadjusted versus Adjusted Six-Year Graduation Rates at Four-Year Colleges and Universities in Virginia



Source: Authors' analysis of Virginia Longitudinal Data System data.

These results mean that state policymakers should look beyond commonly used unadjusted metrics to assess institutional quality. This is especially true when outcomes vary significantly across subgroups or when student characteristics vary significantly across colleges.

We also find that these adjustments cannot be approximated by using only publicly available institution-level data. State policymakers should continue to develop and maintain student-level data systems so these more robust measures can be calculated, allowing them to assess which institutions are best serving their students rather than simply enrolling students who are likely to graduate anywhere. These measures should be straightforward for data analysts working for the state government to calculate, but states would need sufficient resources to employ such staff.

Lesson 2: Completion Rates of Individual Programs Are Important to Measure, but This Can Be Done Accurately Only at Two-Year Institutions.

State and federal policymakers often publish outcome measures (e.g., student debt and postcollege earnings) that drill down within institutions to the program level. But program-level earnings data reflect earnings only for students who graduated with the degree and would ideally be paired with an indication of a student's likelihood of success in a given major within the institution.

This may sound straightforward, but developing program-level graduation metrics is fraught with potholes, particularly for four-year institutions. Using data from Virginia, we identify barriers to producing program-level completion metrics that include as many students as possible, provide an accurate and stable estimate of a student's likelihood of completion, and are consistent across institutions and aligned with institution-level graduation rates.

First, many programs are too small to produce measures of completion that are consistent from one cohort to the next. States can mitigate this drawback by grouping small programs with larger ones (e.g., grouping art subfields together) or by combining multiple years of data, as the federal College Scorecard does by combining two cohorts of entering students. Program-level graduation rates, particularly in small schools, could require additional cohorts of data. The more these data are pooled, the more accurate the metric may be for a typical student's chances of graduating from that program, but the longer it will take for institutions to improve outcomes that are averaged over several years.

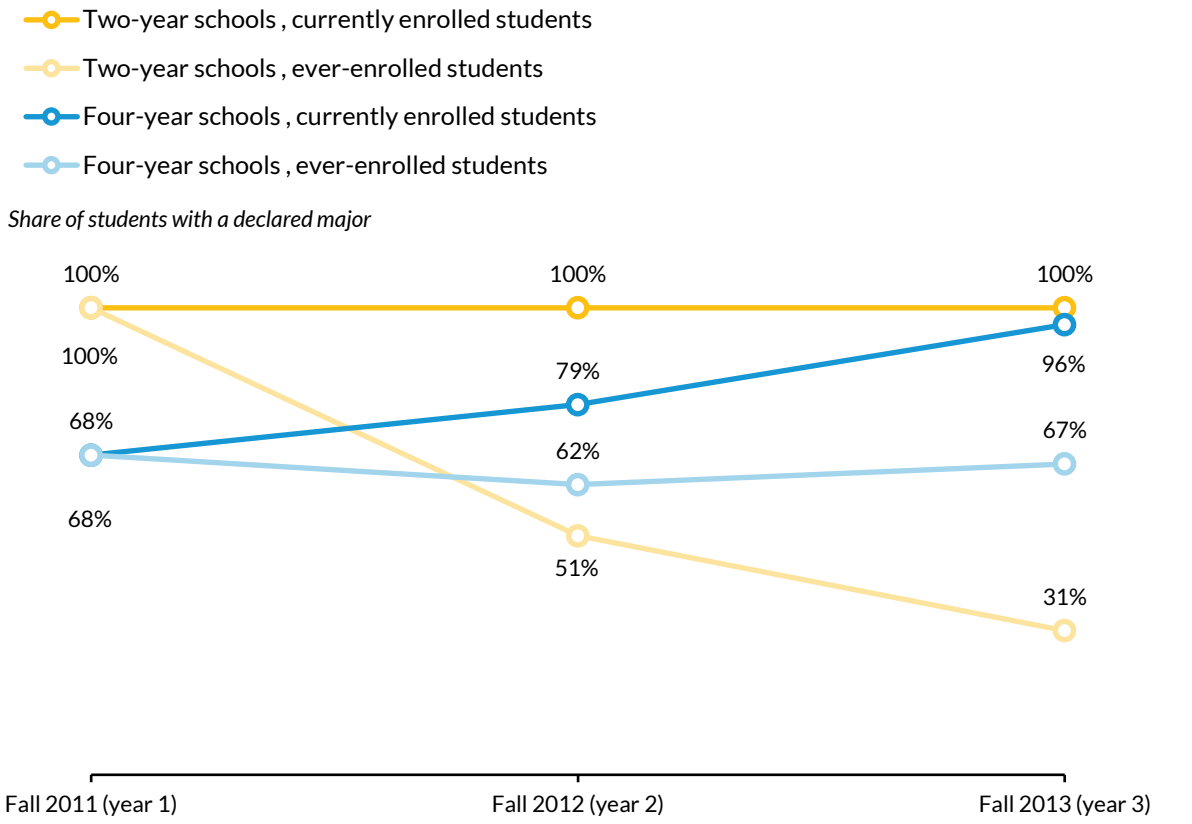
Second, program-level completion can be measured only after students have selected a program. At two-year colleges in Virginia, this is straightforward because all students choose a program upon entry. But at four-year colleges, many students do not declare a major until their second or third year, after some students have already dropped out. Figure 2 displays this tension. In later years, more students have declared majors ("currently enrolled students"), but more have also dropped out (compare with "ever-enrolled students"). Developing accurate program-level graduation rates would have to come with a mandate for students at four-year schools to declare a major by the fall of their sophomore year, at the latest.

For some four-year schools, this could be a large shift, and the cost of mandating this change must be weighed against the potential gain from having these metrics. One potential midlevel step would be

to require students to opt in to a kind of "metamajor"—a large group of potential majors, grouped by subject—similar to what is required of freshman at Georgia State University. This metamajor could be a program-level metric, allowing students the flexibility to select a more specific major later on.

FIGURE 2

Share of Students with a Declared Major, by College Level and Enrollment Status



URBAN INSTITUTE

Source: Authors' analysis of Virginia Longitudinal Data System data.

Third, policymakers need to make a clear distinction between the graduation rate in the same program and the graduation rate in any program in the same institution. To align with earnings data, a within-program graduation rate makes the most sense (as earnings data are for students who earned a degree in the major). But this rate may not reflect variations in the success of students who did not complete a degree in the major but did complete a degree in another major. For example, a student who leaves a math and statistics program and enrolls in, and graduates from, an engineering program would likely be considered a positive outcome, even though she is not counted in the math program's graduation rate. States could provide a more complete picture by publishing both metrics.

Lesson 3: Earnings Are a Popular Outcome Measure, but How They Are Measured Matters.

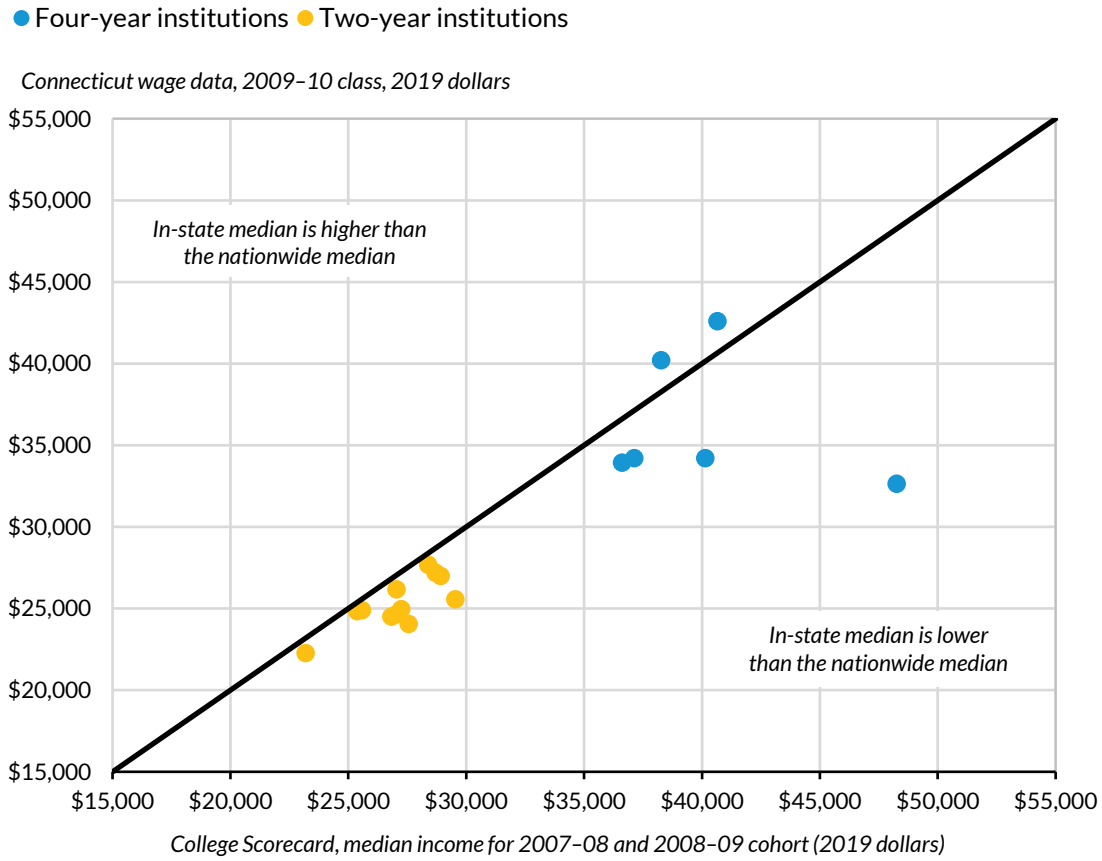
Data on postcollege earnings can help policymakers discern whether attending and completing college increases wages for the state's workers. But data from Connecticut reveal several reasons policymakers should be cautious when using earnings data to compare colleges.

First, commonly used data sources that cannot follow students when they leave the state produce lower measures of earnings than national estimates from the College Scorecard. This is especially true for colleges located near state borders. State policymakers should foster collaborations with federal agencies, such as the Census Bureau's experimental Post-Secondary Employment Outcomes effort, to connect state and federal data.[5]

FIGURE 3

Nationwide versus In-State Median Wages in Connecticut

By level of institution



URBAN INSTITUTE

Source: Authors' analysis of Connecticut Preschool through Twenty and Workforce Information Network and College Scorecard data.

In-state median wages tend to be lower than national median wages (figure 3), and this difference can vary by geography (e.g., how far a college is from a state border). Further, policymakers must consider which cohort provides the most relevant measure. Even though the wages of all those who leave an institution in a given year may provide a more accurate estimate of typical student outcomes, median wage data among graduates tends to more clearly show differences between institutions.

Second, it matters which students are included in the earnings measure and how much time has passed since leaving college. Students who graduate from an institution typically have higher earnings than all of those who leave in the same year, and median wages by institution become more dispersed as former students spend more time in the labor market. Policymakers should find ways to look at multiple measures of earnings, such as an index that reflects both initial earnings and earnings over time.

Third, commonly used earnings measures can exclude the outcomes of many students. Nearly two-thirds of students in a cohort have wage data records in Connecticut, but restrictions that account for other higher education enrollment substantially reduce the number of students included in the measure. Policymakers should describe the students that are included in published earnings measures. There is no one correct approach, so the best solution may be for states to show how they are building their earnings cohort and to allow data users to look at different cohort specifications.

Lesson 4: Data Can Help Policymakers Diagnose the Sources of Equity Gaps in Higher Education, Including the Role Individual Colleges Play and Differences in College Readiness, Financial Circumstances, and Access to High-Quality Institutions.

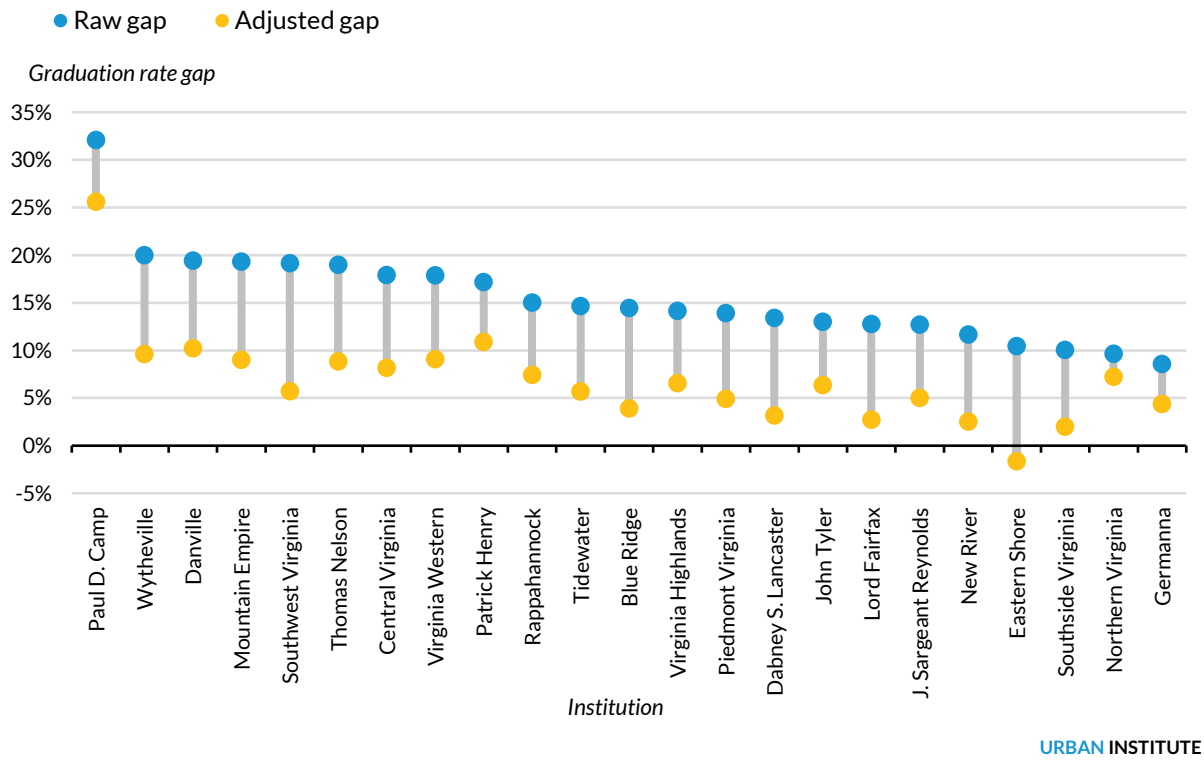
State higher education policymakers raised equity issues in both Connecticut and Virginia, states that have troubling disparities in outcomes between minority students and white and Asian students. Using data from both states, we show how policymakers need to understand how college readiness, individual colleges, and segregation affect outcomes.

First, we show how graduation differences by race or ethnicity shrink when we adjust for such factors as academic preparation and family circumstances, which explain a great deal, but not all, of the gaps in graduation rates (figure 4). Even after adjustments, the gaps at some colleges remain as high as 20 percentage points. The adjusted gaps vary widely, suggesting that some colleges may not support their minority students as much as they support their white and Asian students.

FIGURE 4

Graduation Rate Gaps between Black and Hispanic Students and White and Asian Students

Two-year colleges in Virginia



Source: Authors' analysis of Virginia Longitudinal Data System data.

Second, given the historical relevance of black-white inequities in higher education, we decompose the college graduation gap between these two groups into components associated with racial and ethnic differences in average student characteristics (e.g., SAT scores and high school grade point average), how strongly these characteristics predict graduation, and college quality.

This approach is useful for state policymakers concerned with equity gaps. For example, it shows that equalizing the average difference in SAT scores and high school grade point average by race or ethnicity would close the nearly 20 percentage-point gap in graduation rates at four-year colleges in Virginia by 45 percent. Differences in student financial circumstances explain another 16 percent of the gap. In other words, many of the racial and ethnic disparities at the college level stem from inequality in college readiness and financial strain.

The decomposition also highlights the importance of racial and ethnic segregation between colleges, which could be influenced by state policy. Our estimates suggest that, were black and white students to attend the state's different colleges at similar rates, and black students were to receive the same benefit from attending these schools as white students do, the graduation gap would close an additional 29 percent. This result stems from the fact that even after controlling for SAT scores and high school GPA, black students are less likely to attend high-graduation-rate colleges than white students.

Collectively, college readiness, financial strain, and segregation across colleges explain about 90 percent of the racial and ethnic gap in graduation rates at four-year colleges in Virginia. We find similar results for four-year colleges in Connecticut. Among two-year colleges in both Virginia and Connecticut, differences in full-time status are important components of the graduation gap, while segregation across colleges does not appear to be a factor.

This analysis does not directly identify solutions, but policymakers can use it as a diagnostic tool as they address the roots of inequity in higher education, including income inequality, differential access to K–12 educational opportunities, and unequal access to the colleges that do the best job graduating students.

Conclusion

State higher education policymakers include legislators and legislative staff, officials in governors' offices, education agency staff, higher education coordinating boards, and college administrators. Such a diverse group is sure to have a wide range of needs when it comes to higher education data. And those needs are likely to depend on the state policy and political context, with some states focusing more on economic development and others focusing more on socioeconomic mobility.

State education data systems can be the basis for college performance measures that support informed decisionmaking, whatever the goal may be. Tracking students from high school into higher education and the workforce enables fine-grained analyses that, while not perfect, are more appropriate for informing policy than commonly used measures such as raw graduation rates.

Our analyses of data from Connecticut and Virginia demonstrate the value of these data but only scratch the surface of what is possible. We still have a lot to learn, such as how to adjust or contextualize earnings data based on student characteristics and how to combine sometimes conflicting measures in a way that is useful. We hope these initial efforts will inform researchers working to better measure college performance and state education agencies seeking to use these approaches to produce actionable information for policymakers.

Notes

[1] "10 Essential Elements of Statewide Longitudinal Data Systems," Data Quality Campaign, accessed January 3, 2020, http://dataqualitycampaign.org/why-education-data/state-progress/.

[2] See the website for the US Department of Education's College Scorecard at https://collegescorecard.ed.gov/. See also "Which Colleges in America Help the Most Children Climb the Income Ladder?" The Equality of Opportunity Project, accessed January 3, 2020, http://www.equality-of-opportunity.org/college/.

[3] One recent study found that the release of the earnings data in the College Scorecard had a modest impact on student interest in institutions with high average earnings but only among well-resourced students and schools. See Hurwitz and Smith (2018).

[4] "General Information: Data," Colorado Department of Higher Education, accessed January 3, 2020, http://highered.colorado.gov/data/collection.html.

[5] Post-Secondary Employment Outcomes data provide earnings and employment outcomes by degree, program, and institution for graduates of certain institutions in Colorado, Michigan, Texas, and Wisconsin. These outcomes are tracked nationally, so students are included even if they leave the state after college. See "Post-Secondary Employment Outcomes (PSEO)," US Census Bureau, accessed January 3, 2020, https://lehd.ces.census.gov/data/pseo_experimental.html.

References

Blagg, Kristin, and Matthew M. Chingos. 2016. *Choice Deserts: How Geography Limits the Potential Impact of Earnings Data on Higher Education*. Washington, DC: Urban Institute.

Hurwitz, Michael, and Jonathan Smith. 2018. "Student Responsiveness to Earnings Data in the College Scorecard." *Economic Inquiry* 56 (2): 1220–43.

About the Authors

Erica Blom is a research associate in the Center on Education Data and Policy at the Urban Institute, where she studies higher education policy. Before joining Urban, she studied discriminatory practices in the labor market and anticompetitive practices at Edgeworth Economics. Blom received a BA in mathematics and political science from Queen's University and a master's degree in economics from Western University. She also earned a PhD in economics from Yale University, where her research focused on students' choices in college major. She has published in the Annual Review of Economics, and her dissertation research has contributed to joint work with Benjamin Keys and Brian Cadena on the effects of the business cycle on college major choice.

Kristin Blagg is a research associate in the Center on Education Data and Policy. Her research focuses on K–12 and postsecondary education. Blagg has conducted studies on student transportation and school choice, student loans, and the role of information in higher education. Blagg spent four years as a math teacher in New Orleans and New York City. In addition to her work at Urban, she is pursuing a PhD in public policy and public administration at the George Washington University. Blagg holds a BA in government from Harvard University, an MSEd from Hunter College, and an MPP from Georgetown University.

Matthew M. Chingos directs the Center on Education Data and Policy. He leads a team of scholars who undertake policy-relevant research on issues from prekindergarten through postsecondary education and create tools such as Urban's Education Data Portal. Chingos is coauthor of *Game of Loans: The Rhetoric and Reality of Student Debt* and *Crossing the Finish Line: Completing College at America's Public Universities*. He has testified before Congress, and his work has been featured in media outlets such as the *New York Times*, the *Washington Post*, and NPR. Before joining Urban, Chingos was a senior fellow at the Brookings Institution. He received a BA in government and economics and a PhD in government from Harvard University.

Tomas Monarrez is a research associate in the Center on Education Data and Policy. His research focuses on education policy topics as they relate to economic and racial inequality. Current research projects examine the relationship between school attendance boundary policy and the racial segregation of schools, the impact of for-profit college openings on the structure of local higher education, and the effect of state tuition equity policies on the postsecondary attainment of undocumented immigrant high school students. Monarrez received bachelor's degrees in economics and mathematics from the University of Texas at Austin and earned his doctoral degree in economics from the University of California, Berkeley.

Macy Rainer is a research assistant in the Center on Education Data and Policy, where she focuses on topics in K–12 and higher education. She works on projects related to measures of student poverty, school quality, and college completion rates. Previously, Rainer was a constituent advocate in the office of Boston mayor Marty Walsh. She graduated from Northeastern University with a bachelor's degree in economics.

Kelia Washington is a research analyst in the Center for Education Data and Policy. She supports research centered around K–12 and postsecondary policy issues. Before coming to Urban, Washington worked for the Center for Education Design, Evaluation, and Research at the University of Michigan. She interned at The Century Foundation on their Higher Education Policy team. She also served as a college adviser in rural Virginia with the Virginia College Advising Corps. Washington received her master's in higher education and master's in public policy from the University of Michigan and her bachelor's from the University of Virginia.

Acknowledgments

This brief was supported by Arnold Ventures. We are grateful to them and to all our funders, who make it possible for Urban to advance its mission.

The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of Urban experts. Further information on the Urban Institute's funding principles is available at urban.org/fundingprinciples.

The authors wish to thank David Hinson for copyediting and the participants at convenings held in Hartford, Connecticut, and Richmond, Virginia, in 2018 and 2019, particularly Jan Kiehne and Tod Massa.

500 L'Enfant Plaza SW
Washington, DC 20024

www.urban.org

ABOUT THE URBAN INSTITUTE

The nonprofit Urban Institute is a leading research organization dedicated to developing evidence-based insights that improve people's lives and strengthen communities. For 50 years, Urban has been the trusted source for rigorous analysis of complex social and economic issues; strategic advice to policymakers, philanthropists, and practitioners; and new, promising ideas that expand opportunities for all. Our work inspires effective decisions that advance fairness and enhance the well-being of people and places.