IES Institute of Education Sciences

Steps to Develop a Model to Estimate School- and District-Level Postsecondary Success

Regional Educational Laboratory Central

At Marzano Research

REL 2021–074
U.S. DEPARTMENT OF EDUCATION

A Publication of the National Center for Education Evaluation and Regional Assistance at IES

Steps to Develop a Model to Estimate School- and District-Level Postsecondary Success

R. Marc Brodersen, Douglas Gagnon, Jing Liu, and Tony Moss

March 2021

This tool is intended to support state and local education agencies in developing a statistical model for estimating student postsecondary success at the school or district level. The tool guides education agency researchers, analysts, and decisionmakers through options to consider when developing their own model. The resulting model generates an indicator of a school's or district's contribution to the postsecondary success of its students after contextual factors are accounted for that might be outside a school's or district's control, such as student demographic characteristics and community characteristics. State and local education agencies could use the information generated by the models they develop to help meet federal and state reporting requirements and to inform their own efforts to improve their students' postsecondary success.

Contents

About this tool

This tool describes a step-by-step process for developing a statistical model to estimate student postsecondary success at the school or district level—that is, the accomplishments of students after they leave high school for college, the workforce, the military, or civic life. The tool draws on support that the Regional Educational Laboratory Central provided to the Kansas State Department of Education to review and refine the Kansas Postsecondary Effectiveness Model. Although the considerations presented for developing a postsecondary success model can apply to districts as well as schools, for ease of exposition this tool refers mainly to schools.

Postsecondary success models are designed to estimate the influence that schools have on their students' success after high school, after contextual factors are accounted for that are partially or wholly outside schools' control. These models use data such as postsecondary enrollment or workforce employment that can reveal how successful students are after high school. To account for contextual factors, these models also use data collected while students are still in high school, such as demographic and community characteristics.

This tool is designed to help education leaders, including researchers and data analysts, who are responsible for implementing or supporting efforts to promote students' postsecondary readiness and success goals in their schools. Education leaders can use the information generated by their postsecondary success model to help meet federal reporting requirements and to enable their schools to compare their performance with that of other schools. Education leaders can also use the data generated by the model to identify schools that are struggling to prepare students for college or careers, which might warrant further inquiry into root causes and could yield information for making programmatic decisions or targeting extra support. Additionally, to support improvement strategies, education leaders might want to pair schools that have consistently low estimates of postsecondary success with similar schools that perform well. Finally, education leaders might want to acknowledge schools that consistently perform well on indicators of postsecondary success by designating them as high-performing schools and inviting school representatives to speak about their strategies at state, regional, local, and other events.

Although the intended audience for this tool has some experience with statistics, key terms are defined in box 1 as a reminder.

Box 1. Key terms

Contextual factors. Student demographic characteristics and community characteristics that are partially or wholly outside schools' control. These characteristics may be included as covariates in post-secondary success models.

Correlation. A measure of the strength of the relationship between two variables. Household income and student eligibility for the national school lunch program would be expected to have a high correlation because they are both proxies for socioeconomic status. However, knowing that two variables are correlated is not sufficient to claim that one causes the other.

Covariate. A variable that has a relationship with the dependent variable (student postsecondary success) and that can be accounted for when examining a school's influence on student postsecondary success. Covariates included in the model are contextual factors such as the percentage of students eligible for the national school lunch program, a proxy for low-socioeconomic status.

Cutscore. Values of a postsecondary success indicator that can be used to classify schools into performance categories. For example, a cutscore might be set to distinguish between schools classified as meeting expectations and schools classified as excelling.

Pooling data. Aggregating data across students who have different expected high school graduation years.

Postsecondary success. Students' accomplishments after high school—in college, the workforce, the military, or civic life.

Postsecondary success indicator. An estimate of a school's influence on student postsecondary outcomes, after contextual factors are accounted for. The indicator is an estimate of the extent to which the school is doing better or worse than expected, given the influence of the contextual factors.

Postsecondary success score. The number or percentage of students in a school who meet the definition of postsecondary success used in the postsecondary success model.

Regression. A statistical procedure used to examine the relationship of one or more variables to student postsecondary success.

Regression coefficient. A statistical estimate of the relationship between a given variable and student postsecondary success, after the relationships between the other variables and student postsecondary success are accounted for.

School postsecondary success. The extent to which a school's students demonstrate postsecondary success.

Student cohort. A group of students expected to graduate from high school in the same year.

Student postsecondary success. The extent to which students achieve success in their life after high school, defined in terms of college, workforce, military, or civic outcomes.

Overview of the development of the postsecondary success model

The postsecondary success model described in this tool estimates a school's influence on student postsecondary success after contextual factors are accounted for that are wholly or partially outside the school's control, such as student demographic characteristics and community characteristics. The model generates a postsecondary success indicator that is an estimate of the extent to which the school is doing better or worse than expected on student postsecondary success, given the influence of the contextual factors. By accounting for these contextual factors, model estimation can provide a more accurate representation of a school's influence on student postsecondary success than can a simple examination of the percentage of a school's students who demonstrate success on various postsecondary outcomes.

The postsecondary success models described in this tool are similar to other value-added models, such as models that estimate teachers' effect on student achievement by accounting for factors outside teachers' control and models that use student data to predict the risk for dropping out of school. For example, teacher value-added models might account for students' prior academic achievement and demographic characteristics to estimate teachers' effect on their students' current academic achievement. The primary difference between these types of models and postsecondary success models is that teacher value-added models attempt to attribute effects to individual teachers, whereas postsecondary success models attempt to attribute effects to schools or districts.

This tool outlines a process for developing a model to generate a postsecondary success indicator for schools. The considerations described in the tool are intended to support education agency researchers, analysts, and decisionmakers in developing a model that accounts for contextual factors that might influence student postsecondary success. The types of models described in this tool can be developed using data from statewide longitudinal data systems, which are available to most states, and from other accessible data sources such as the National Student Clearinghouse.

This tool describes five steps to consider when constructing a postsecondary success model:

1. Determine what data are available to measure postsecondary success.
2. Select covariates for the model.
3. Develop the model.
4. Use postsecondary success indicators.
5. Compare different models.

Before developing your model, consider the questions in box 2, which guide you through the model development steps (box 2). It might be helpful to refer to the answers to these questions as you develop your model. How the model results will be interpreted and used has direct implications for the validity of the model. For instance, you might determine that the primary problem you want to solve is how to make higher education more affordable. In that case you might decide not to develop a postsecondary model if you determine that the model cannot be used to support increased college success. Alternatively, you might decide that a postsecondary model makes sense in your context. If the model will be used for high-stakes purposes, consider additional factors related to validity and unintended consequences. These considerations are discussed throughout this tool.

Box 2. Guiding questions while developing the postsecondary success model

Before developing a postsecondary success model, ask why you want the model and how you intend to use it. Because you might have to justify the selection of various model components and explain how the results will be used, your answers to these questions should guide all aspects of model development.

Questions to consider include the following:

- What problem are you trying to solve?
- Why is measuring postsecondary success important in your state or district?
- In what ways can the model address strategic priorities for your state or district?
- What contextual factors might influence student postsecondary success in your state or district?
- How do you intend to use results generated by the model? Will high-stakes school accountability decisions such as funding be informed by model outputs?

The next section addresses key considerations for each model development step (figure 1 presents an overview). Step 1 is of interest for selecting and constructing measures of student postsecondary success. Steps 2 and 3 guide you through the process of developing your own postsecondary success model. Finally, steps 4 and 5 provide recommendations for how to use the results from your postsecondary success model and how to compare the results generated by different model types. For an example of a completed model, see the detailed description of the Kansas Postsecondary Effectiveness Model in appendix A.

Figure 1. Steps and key considerations in developing a statistical model for assessing student postsecondary success

Model development step

Key considerations

↓

Step 1: Determine what data are available to measure postsecondary success

Which postsecondary success domains and measures will you include?	What data are available?	How many measures will you use?

↓

Step 2: Select covariates for the model

What contextual factors should you account for?	How will you measure covariates that change over time?	How do covariates influence the ease of model interpretation?

↓

Step 3: Develop the model

Will your model use a single-level or mutilevel design?	Will you use a fixed-effects or a residual approach?	Will you pool student data?

↓

Step 4: Use postsecondary success indicators

How will you classify schools?	Will you set cutscores?

↓

Step 5: Compare different models

How sensitive are postsecondary success indicators to model variations?	How stable are postsecondary success indicators?	How do postsecondary success indicators vary across school type?

Source: Authors' construction.

Steps in developing a postsecondary success model

Step 1: Determine what data are available to measure postsecondary success

In developing a postsecondary success model, first determine how you want to define and measure student postsecondary success and identify what data are available or could be easily obtained. You might also consider whether to assess postsecondary success over the short or long term (for example, at six months or at two years after high school) and whether to use single or multiple indicators of postsecondary success.

Which postsecondary success domains and measures will you include?

Postsecondary success can be defined in relation to four common domains: college, workforce, military, and civic life (Adelman, 2006; Conley, 2014; Gheen et al., 2012). For each domain there are a variety of ways to measure success (table 1). You might want to select a definition of postsecondary success that aligns with state or district priorities. For example, if a state education agency has set supporting career pathways as a strategic priority, you might consider incorporating workforce measures into your definition.

Table 1. Example of postsecondary success domains and measures

College	Workforce	Military	Civic
Enrollment immediately after high school	Entry into the workforce immediately after high school	Armed Services Vocational Aptitude Battery scores	Civic attitude (for example, sense of responsibility to the community)
Two-year continuous enrollment	Employment status	Military enlistment	Political attitude (for example, sense of responsibility to engage in the political process)
Credit accumulation	Continuous employment		Civic participation (for example, volunteering)
Transferring from two-year to four-year institution	Quarterly wages		Political participation (for example, voting, protesting)
Degree attainment	Professional certification		

Source: Authors' compilation.

What data are available?

Another important consideration when developing a postsecondary success model is to determine what data are currently available or could be easily obtained. Your postsecondary success model will need some measure of student success that occurs after high school. When deciding how to define and measure student postsecondary success, begin by inventorying your data system to identify relevant data that are currently being collected by your education agency or that can be accessed through agencies with which you have data-sharing agreements. For example, your agency might already store student demographic data in its longitudinal data system or collect postsecondary education information through the National Student Clearinghouse.

Alternatively, you could select a definition of postsecondary success that requires collecting new data. For instance, including civic readiness as part of the definition of postsecondary success might require developing a survey instrument to capture the perceptions of high school graduates on civic matters. Similarly, including employment or wage data might require developing a data-sharing agreement with the state department of labor or another agency. Although collecting new data allows you to expand your conception of postsecondary success, it also creates a new burden for leaders and educators.

Postsecondary success can be measured in the short term (for example, two years after high school) or the long term (for example, five years after high school). In determining at what point to measure student postsecondary success, consider how much time is typically required to achieve the desired postsecondary outcome (for example, immediate entry into the workforce or four years to complete a college degree), whether the relevant data are available within that timeline, the likelihood that more data will be missing as the timeline lengthens, and the diminishing influence of secondary education factors on more distant outcomes.

Missing data are more of a concern if you pursue longer-term outcome measures. Linking secondary school data to college or workforce data typically becomes more difficult the longer students are out of high school—for instance, as individuals move out of the state or change their name. Perhaps an even more important concern than missing data is the greater difficulty of attributing more distant outcomes to secondary school influence. Thus, although longer-term outcomes might be more meaningful than short-term outcomes (for example, college graduation as opposed to college enrollment), the link between student outcomes and the high schools they attended becomes weaker as one looks further out from high school graduation.

How many measures will you use?

Another important decision is whether to use a single measure or multiple measures of postsecondary success. Using a single measure (for example, quarterly wages) supports a more straightforward model, limits the amount of data needed, and eases interpretation of the results. However, there are drawbacks to defining postsecondary success too narrowly, as students can take any of multiple pathways to success after high school. You might wish to capture these different pathways by incorporating measures across multiple success domains. You can use multiple measures of student postsecondary success separately or combine data across multiple domains and measures into a single postsecondary success score or index (Conley, 2014). For example, a school might get a point for each student who enrolls in college, attains a professional certification, or maintains steady employment for a year. Alternatively, schools could be rated separately on their graduates' success in college and in the workplace. Although this approach incorporates multiple definitions of success, it requires additional data that likely need to be gathered from different agencies. Examples of how two state models incorporate multiple measures of postsecondary success are in box 3.

Box 3. Examples of postsecondary success measures in two state models

Kansas Postsecondary Effectiveness Model: Kansas's model combines two measures—the percentage of students from a district who are continuously enrolled in a postsecondary institution for two years after their expected high school graduation year and the percentage who earned a postsecondary credential within two years of their expected graduation year—into a single postsecondary success score (see appendix A for more details about this model).

Louisiana Promotion Power Model: Louisiana uses five measures of postsecondary success and runs the model separately for each (Deutsch et al., 2020):

- On-time high school graduation.
- Completion of a college or career readiness credential (completing an Advanced Placement, International Baccalaureate, or dual enrollment class or demonstrating proficiency in a recognized skill area).
- College enrollment.
- Multiyear college persistence (attending college for at least four of the five years after the expected high school graduation year).
- Earnings at age 26.

Step 2: Select covariates for the model

The next step in developing a model is to decide which covariates to include to account for contextual factors that could influence student postsecondary success but that are partially or wholly outside schools' control, such as student socioeconomic status. The rationale is to ensure that schools are not unfairly benefited or penalized by such contextual factors.

To identify an initial list of covariates, first examine prior research. For example, Hein et al. (2013) provide an overview of student skills and behaviors that have been linked to postsecondary success. Additional factors to consider include whether the covariates can be quantified, the relationships between covariates and postsecondary success rates, and whether the covariates make it easier or harder to interpret the model results.

What contextual factors will you account for?

In selecting covariates to include in the model, consider the extent to which schools are held accountable for student-level, family-level, and community-level characteristics that are likely to affect student postsecondary success. Including covariates conveys that they are partially or wholly outside schools' control. You should again consider the intended purpose of your model and how the model can address state or district priorities.

For example, a state education agency in a rural state might develop a definition of postsecondary success that includes rates of college enrollment. This agency might decide to account for distance to the nearest college in its model so as not to penalize rural schools that are farther away from colleges, as students in rural areas might be less likely to attend college (National Student Clearinghouse, 2016).

These considerations apply to a host of factors that are linked to student postsecondary success, including student factors, such as disability or chronic absenteeism, and community characteristics, such as economic health and distance to a college. However, do not include covariates over which schools have more control, such as teacher evaluation scores, dual enrollment participation, or number of counselors in a school. These covariates are likely related to a school's influence on student postsecondary success, so including them in your model would account for some of the what the model is designed to measure. Therefore, carefully consider which covariates to include and be able to justify their inclusion.

Implications of covariate selection

Accounting for certain factors based on problematic assumptions could lead to perverse incentives. For instance, accounting for the percentage of English learner students in a postsecondary success model might lead schools to classify more students as English learner students. Additionally, accounting for student subgroups could be interpreted as tacitly suggesting that schools are less responsible for the outcomes of these subgroups and that the bar for success is lower for these students. Moreover, although schools cannot control the populations they serve, they can intervene in ways that might support student postsecondary success.

However, factors should be considered for inclusion if they can help place schools on a more equal level. For example, lower-performing students are much less likely to attend college (Caro et al., 2015). However, high schools have little control over the previous achievement of students. Schools that serve higher percentages of lower-performing incoming students, therefore, would be at a disadvantage if their postsecondary success were to be estimated without taking this covariate into account.

How will you measure covariates that change over time?

After identifying a set of possible covariates, determine how to turn the covariates into measurable factors. One consideration is whether to use cross-sectional or longitudinal student data. Additional statistical considerations for selecting covariates are discussed in box 4.

Cross-sectional data are gathered for each school or each school's relevant graduating student cohorts at a single point in time, such as grade 12. Although these data might be easy to access and use, they do not account for the possibility that student status in some demographic categories might change over time (for example, English learner status or eligibility for the national school lunch program).

Alternatively, you might take advantage of the student-level longitudinal data systems available to most state education agencies. You might use a multilevel model (see step 3) or calculate the average percentage of time during which students in a school were part of a particular demographic category. Accounting for possible variation over time in some student characteristics could make the model more accurate. However, this approach might require more analytic capacity, and the model could be difficult to describe to others. The steps for calculating aggregated longitudinal covariates are detailed in appendix B.

Box 4. Statistical considerations for selecting covariates

There are several statistical factors to consider when selecting covariates for your postsecondary success model.

Relationship of covariates to postsecondary success. Because postsecondary success models are designed to account for the influence of contextual factors (covariates) on student postsecondary success, covariates should be correlated with student postsecondary success or theoretically be able to impact it.

Improved prediction of postsecondary success. When selecting an additional covariate to include in your model, examine whether it improves the prediction of student postsecondary success beyond the predictive power of covariates already included in the model. Keep in mind that the purpose of including multiple covariates in your model is not simply to maximize the model's explanatory power. Doing so risks including variables that represent internal factors that may be part of the school's influence on student postsecondary success.

Relationships of covariates with one another. Including multiple covariates that are highly correlated with one another makes it more difficult to interpret their individual relationship to postsecondary success. However, this is not a major concern for postsecondary success models because their purpose is to generate school postsecondary success indicators and not to assess the relative influence of covariates on those indicators.

How do covariates influence the ease of model interpretation?

You might also consider how your choice of covariates influences the ease of interpreting and explaining the model's results to others. Generally, simpler models (such as those with fewer variables) are easier to interpret. Thus, if there is not a strong rationale for adding a particular covariate or if your model already includes similar covariates, you might want to exclude the additional covariate to keep the model simpler and make it easier to interpret the results. For instance, to account for a community-level income variable, you might consider both median household income and per capita income as possible covariates in your model. These two variables are likely very closely related, and there may not be strong justification for including both measures. Including only one of the variables keeps the model simpler while still accounting for the underlying community characteristic of interest.

Covariate justification

The intended purpose and use for your model should guide your selection of covariates. If you have a justification or theory for including a particular covariate in a model and you consider it to be an important external factor to account for, its inclusion will not harm your model.

Step 3: Develop the model

You have several options when developing your model to generate postsecondary success indicators. These options include whether to use a single-level or multilevel model and whether to use a fixed-effects or a residual approach for the model (see appendix B for the statistical equations for each model type). For some model alternatives you might incorporate data from a single student cohort or from multiple cohorts. The implications of this choice, including the advantages and disadvantages of pooling data when using multiple cohorts, are discussed below.

Will your model use a single-level or multilevel design?

As you develop your model, you have the option to use a single-level or multilevel design. An example of how the data associated with each design would be organized is in table 2.

Single-level design. Single-level models use data from a single cohort of students or pooled data from multiple cohorts. In each case there is a single data point for each variable for each school. This model will generate a single postsecondary success indicator for each school.

Multilevel design. Alternatively, you could use a multilevel model in which data from multiple student cohorts are nested within schools.[1] For example, you could incorporate data from three consecutive graduating cohorts. Rather than the data from these three cohorts being pooled, as in a single-level model, each cohort has its own data point for each variable, and data for each cohort are associated with the relevant school. This approach helps account for the variability of data across each school's cohorts. It also accounts for the dependence of cohort outcomes within a school over time since cohorts within a given school are likely to be more similar to one another than to cohorts in other schools. A multilevel model will generate a separate postsecondary success indicator for each school's cohorts included in the model.

Although the multilevel design adds precision, it requires greater analytic capacity and can be more difficult to describe to stakeholders. Because this approach generates multiple postsecondary success indicators for each school, leaders will need to determine how to interpret and use these indicators. For example, leaders could pool the multiple indicators for each school or examine the most recent indicator (these options are discussed in the section on pooling data below).

1. You could also include additional levels of data in your model, such as time nested in students, students nested in cohorts, and cohorts nested in schools or districts (Raudenbush & Bryk, 2011). However, discussion of these more complex models is beyond the scope of this report.

Table 2. Single-level designs have one line of data for each school, whereas multilevel designs have one line of data for each student cohort in each school

Single-level design				Multilevel design				
School ID	Covariate 1	Covariate 2	Student postsecondary success	School ID	Student cohort	Covariate 1	Covariate 2	Student postsecondary success
School A	87	47	86	School A	2017	87	40	92
School B	36	83	75	School A	2016	91	46	85
School C	45	55	94	School A	2015	85	52	88
School D	72	68	67	School B	2017	42	81	70
School E	58	39	81	School B	2016	33	94	74
School F	66	45	50	School B	2015	39	91	78

Source: Authors' construction.

Additional details on both approaches are in appendix B.

Will you use a fixed-effects or a residual approach?

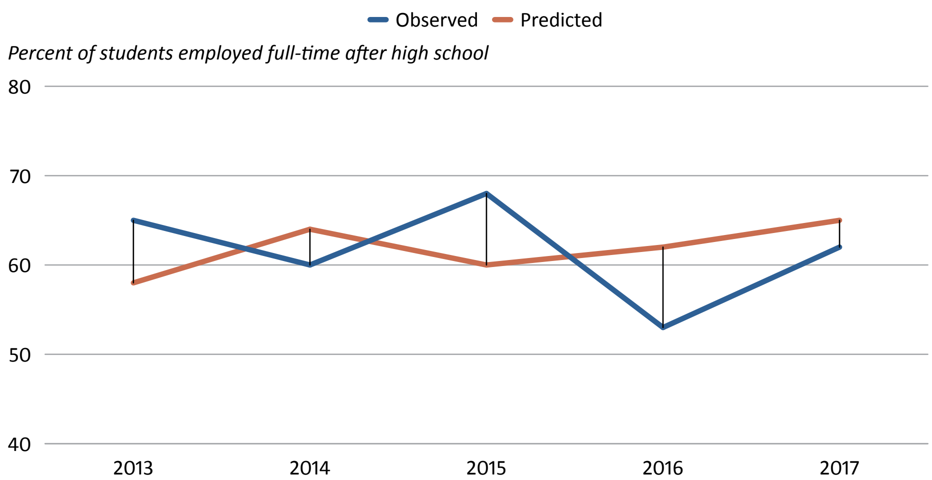
Also determine whether to use a fixed-effects or a residual approach to generate post-secondary success indicators (see appendix B for more technical descriptions).

Fixed-effects approach. A fixed-effects approach uses a separate variable to represent each school included in the model. A school's postsecondary success indicator is the regression coefficient associated with that school's variable. The regression coefficient for each school is generated after accounting for the influence of the model's covariates on student post-secondary success. The larger the coefficient, the larger the school's estimated association with student postsecondary success.

Residual approach. A residual approach uses a school's covariate values to generate a predicted student postsecondary success score for each school instead of including a separate variable for each school in the model. A school's postsecondary success indicator is the difference, called a residual score, between its predicted and actual or observed post-secondary success scores (see figure 2 for an example). If the observed score is higher than the predicted score (in other words, the residual is positive), the school performed better than expected given its student or community characteristics.

Currently, there is little guidance on whether to use a fixed-effects or a residual approach. Both approaches require roughly equal analytic capacity and burden. Research on value-added models also suggests that the two will produce results that are highly correlated (Rose et al., 2012). Additionally, the decision to use a fixed-effects or a residual approach has no implications for the number of postsecondary success indicators generated by the model (table 3). Ultimately, this discussion of fixed-effects and residual approaches aims to inform you of the options, but a deeper discussion of their relative merits is technical and beyond the scope of this tool.

Figure 2. In this example of a residual approach, the school's estimated postsecondary success indicator is the difference between the observed and predicted employment success of its students



Note: The difference between a school's observed and predicted student postsecondary success score at each time period, signified by the vertical bars, represents the postsecondary success indicator for the corresponding student cohort in a multilevel model. A single-level model would generate an indicator only for an individual cohort.

Source: Authors' construction.

Table 3. The number and type of postsecondary success indicators generated by a postsecondary success model depend on how the model is specified

Model type	Single-level model	Multilevel model
Fixed-effects model	Single school-level regression coefficient	Multiple cohort-level regression coefficients
Residual model	Single school-level residual score	Multiple cohort-level residual scores

Note: Example statistical equations for each model specification are in appendix B.

Source: Authors' construction.

Will you pool student data?

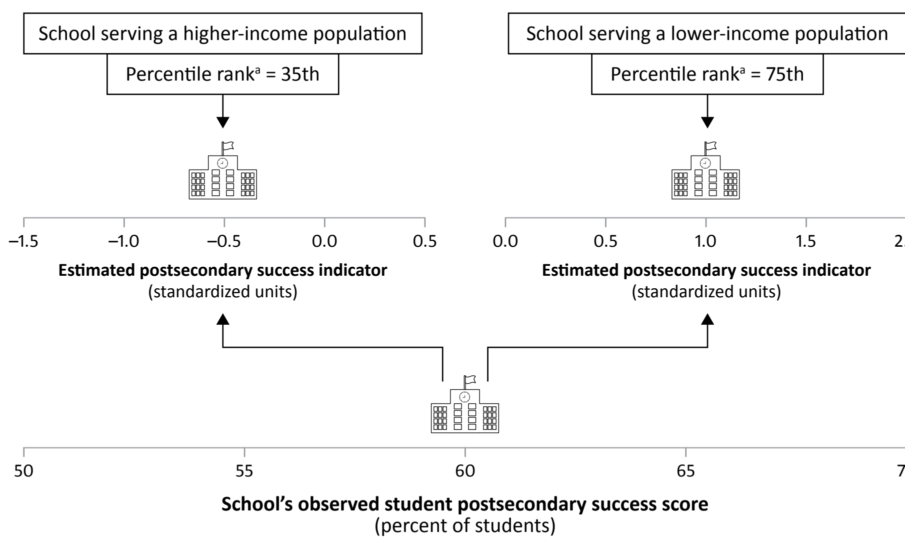
You have the option of using data from a single student cohort or multiple student cohorts in your postsecondary success model. A school-level postsecondary success indicator generated using a single cohort allows you to estimate yearly indicators. This could be desirable, as it theoretically allows you to see how postsecondary success indicators change from one year to the next. However, indicators based on a single cohort of data can be highly imprecise because of idiosyncrasies of that cohort or time period. Any changes could thus be the result of random variation or another underlying phenomenon (for example, regression toward the mean). Consequently, you run the risk of drawing unjustified conclusions from an imprecise model. This is particularly likely in the case of small schools, where the small student population can result in substantial changes in outcomes from year to year.

Alternatively, you could use data from multiple student cohorts when estimating postsecondary success indicators. Pooling data on multiple cohorts of students is more likely to smooth out random errors so that postsecondary success indicators are more stable. This can be done before running the postsecondary success model or by averaging the multiple, single-year postsecondary success indicators for each school afterward. This method allows for more stable indicators, but it has the limitation that performance in past years will continue to affect indicators for some time, regardless of how well the school has performed in the most recent year.

Step 4: Use postsecondary success indicators

Once you have generated postsecondary success indicators for each school, decide how the indicators will be used or interpreted. Postsecondary success indicators for individual schools, whether derived from a fixed-effects or a residual approach, are interpreted in the context of indicators generated for all other schools. Additionally, because covariates are used to account for contextual factors, schools with similar student postsecondary success scores but with different student demographic characteristics or community characteristics could have different estimated postsecondary success indicators. For example, two schools might both have an observed postsecondary success score of 60 percent, but because they serve different student populations (for example, with different average household income) they have different estimated postsecondary success indicators. As a result, the school serving the lower-income population is ranked higher than the school serving the higher-income population (figure 3).

Figure 3. Schools with similar observed student postsecondary success scores might receive different indicators when their student populations differ



Note: The figure shows two schools that have the same observed postsecondary success score of 60 percent. However, the model results in a lower estimated postsecondary success indicator and a lower percentile rank for the school that serves a higher-income population than it does for the school that serves a lower-income population.

a. The school's percentile rank on the postsecondary success indicator.

Source: Authors' construction.

The simplest approach to interpreting postsecondary success indicators, whether they are regression coefficients or residual scores, is to rank order them, placing the schools along a continuum from lowest to highest based on their postsecondary success indicators. Since the models described here generate school postsecondary success indicators in comparison with those of other schools, postsecondary success indicators can be used only to make relative judgments in a zero-sum fashion. In other words, a gain on a postsecondary success indicator from one year to the next in one school must be accompanied by a loss among the remaining schools. Setting strict criteria for determining school effectiveness based on schools' postsecondary success indicators (for example, the top 20 percent of schools) will result in a predetermined percentage of schools receiving effective or ineffective designations over time, even if schools improve or decline in their actual influence on the postsecondary success of students. Ultimately, you will have to determine how you want relative school effectiveness and ineffectiveness to look in your particular context.

How will you classify schools?

If you intend to use postsecondary success indicators as an informal performance indicator, simply providing schools with their indicators and information about where they fall on the distribution could be sufficient. In this case schools would receive information about how their postsecondary success compares to that of other schools. If you want to classify schools relative to one another in some way (for example, as struggling, meeting expectations, or excelling), you might opt to classify schools using percentile rankings or a standard deviation distance away from the average indicator.

Will you set cutscores?

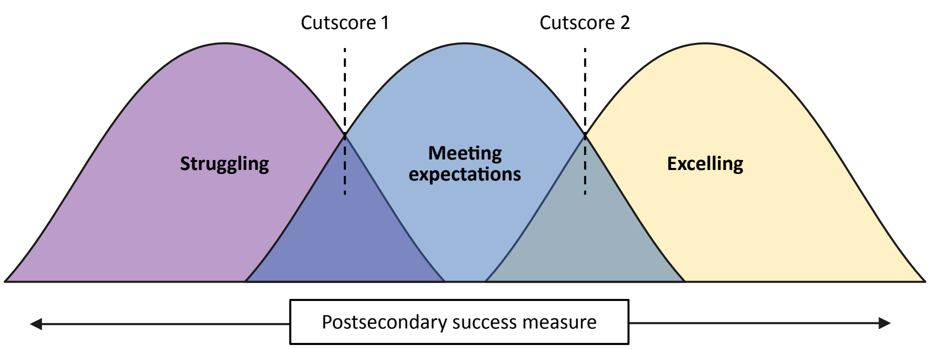
Although postsecondary success indicators provide a way to rank order schools while their different contexts are accounted for, they might be difficult for stakeholders to understand. For a more rigorous process to differentiate schools based on their postsecondary success indicators, you might consider using a standard-setting approach to set performance cutscores that are linked to clear success indicators (Cizek & Bunch, 2007). One example is the contrasting-groups method (Berk, 1976), described briefly here and elaborated in Cizek and Bunch (2007).

Imagine that you have established three performance categories—struggling, meeting expectations, and excelling—with qualitative descriptions illustrating where schools at each performance level tend to do well and where they struggle. The contrasting-groups method can establish the two cutscores separating the struggling and meeting expectations categories and the meeting expectations and excelling categories. You draw a representative sample of schools and convene a group of subject matter experts who are familiar with those schools and with school success metrics more generally. The schools should vary in their position in the rank order of their postsecondary success indicators.

The experts are given the raw data (the covariate and outcome data) for each school. Their task is to assign each school to a performance category based on their knowledge of the school, a holistic appraisal of that school's covariate and outcome data, and the qualitative descriptions of each performance level. The contrasting-groups method does not require the experts to use the same criteria to classify schools. Rather, they examine the body of data for each school and make a holistic judgment of that school's performance given its particular context. If the experts' classification of the schools diverges substantially from the schools' postsecondary success indicators, the validity of the experts' classification or the indicators would be questioned. In this case, both the postsecondary success model specifications and the expert classification process should be examined to identify the cause of the divergence.

To establish the cutscores for the postsecondary success indicators, inspect the three overlapping distributions of postsecondary success indicators for the schools that the experts have placed in each performance category. The cutscores are generally placed at the points where the distributions cross. Alternatively, you can use logistic regression to find the postsecondary success indicators associated with a 0.5 probability of being in one performance category (such as struggling) rather than the others. An example of a postsecondary success indicator distribution is in figure 4.

Figure 4. Subject matter experts' categorizations of schools' postsecondary success can be used to set postsecondary success indicator cutscores



Source: Authors' construction.

Step 5: Compare different models

Given the importance of accurate and fair postsecondary student success indicators and the complex assumptions that must be made in generating them, you should examine the stability of these indicators as you make decisions about how to construct your model (Newton et al., 2010). This section discusses sensitivity analyses that you could conduct to understand how robust postsecondary indicators are with respect to your choice of model.

How sensitive are postsecondary success indicators to model variations?

One family of analyses examines how postsecondary success indicators change with model variations, such as using different postsecondary success indicators or covariates, using a single-level or multilevel model, or including a single student cohort or multiple cohorts.

You might document any particularly contentious or arbitrary decisions that you made while constructing your preferred model. Next, you might look at how the results vary if you make different decisions.

For instance, you might examine the percentage of schools that change education agency–determined effectiveness categories as you make changes to your model, such as including or excluding particular covariates. For example, this would be relevant if you determined the effectiveness categories of schools by categorizing the bottom 10 percent of schools as ineffective, the top 10 percent as highly effective, and those in between as effective.

Alternatively, you could examine the average change in schools' indicator percentiles across different model options. If changing a model option results in a relatively large shift in postsecondary success indicators, that model option is highly consequential to the model's results. For instance, if you are unsure of whether to pool multiple cohorts, you could compare a model without pooling to one that pools five years of data and calculate the average difference in the percentile ranking for schools under the two scenarios. If you find large differences in percentile rank in the two cases, you might opt to use multiple years of data in order to smooth out errors across years, or you might decide to be very cautious in using the postsecondary success model in high-stakes contexts. You can run similar comparisons for other model options as well, such as which covariates to include or whether to use a single-level or a multilevel model.

You might also examine correlation coefficients between postsecondary success indicators generated by different models to see whether various decisions produce similar or dissimilar results (Newton et al., 2010). If indicators differ considerably according to rather arbitrary decisions about model options, you should be very cautious in using your postsecondary success model even in relatively low-stakes contexts. For instance, you might find that including a certain covariate results in 20 percent of schools changing performance categories. Such a large change can be worrisome and might dissuade you from using the postsecondary success model for higher-stakes purposes.

How stable are postsecondary success indicators?

You might want to examine how stable your postsecondary success indicators are from one year to the next. Again, you might want to consider the percentage of schools that change performance categories, or the average percentile point change, across time. If there are large changes in school postsecondary success indicators, on average, from one year to the next, you should be very cautious in using these indicators to make high-stakes decisions. Greater year-to-year fluctuation in student achievement data is common among schools with lower enrollments.

You might also examine how much the postsecondary success indicators of schools change relative to other schools in the state from year to year. If many schools change performance categories or move substantially up or down in the distribution, you might need to pool data across multiple student cohorts to smooth out the variation. However, if the indicators (or the underlying measure of student postsecondary success) are highly unstable over time, aggregating these data might not be meaningful, as your model might ultimately be capturing too much noise and too little signal.

How do postsecondary success indicators vary across school type?

You might also examine how key school characteristics relate to postsecondary success indicators across a range of model options. Possible school characteristics include the percentage of students from low-income households, school locale (city, suburb, town, or rural), enrollment, and racial/ethnic composition. This examination might reveal that some model options (such as accounting for certain covariates or using a single-level rather than a multilevel model) are particularly advantageous or disadvantageous for certain types of schools or groups of students.

For instance, a model that does not account for locale might result in disproportionately low indicators for rural schools. In that case you might account for locale. This is essentially a judgment call because there is no way to know to what extent rural schools are being assessed fairly (for example, because they are not adequately preparing students for the rigors of college) as opposed to being unfairly penalized (for example, because rural schools are located farther from colleges).

Ultimately, objective, agreed-on guidelines are lacking for establishing the validity of a postsecondary success model. Use your judgment to determine which models are stable across model options, time, and school characteristics. Postsecondary success models that will be used for higher-stakes purposes should be highly stable, as it will be difficult to justify decisions based on data from models that result in drastically different outcomes when specifications change.

Limitations

There are limitations to postsecondary success models in general. Most important, these models cannot fully account for contextual factors that are outside schools' control and that affect students' postsecondary success. Unmeasured or unobserved factors might influence whether a given student enrolls in college or otherwise demonstrates postsecondary success. Conversely, factors that are included in postsecondary success models as covariates could, to some degree, be under the influence of schools, and therefore postsecondary success models that account for contextual factors could underestimate the influence of schools. For instance, you might want your model to account for school funding, as this is largely outside schools' control. However, it could be that schools with greater funding are able to provide more support for students' transition to postsecondary life. Ultimately, it is difficult to isolate the influences of a given school on an outcome with so many contributing factors.

The nature of the postsecondary success models described in this tool could make it difficult for some schools to receive very high or very low postsecondary success indicators. The models generate postsecondary success indicators that signal whether a school performed better than expected given its student or community characteristics. For instance, if a model accounts for the percentage of students from low-income households in a school, it could be difficult for high-performing schools with large percentages of students from high-income households to score much better than expected because that expectation is already nearing an upper limit close to 100 percent ("ceiling effect"). Alternatively, it might be difficult for low-performing schools with large percentages of students from low-income households to score worse than expected. Additionally, the indicators generated by the models are most useful for identifying very high- or very low-performing schools. Be cautious in making distinctions between schools with very similar indicators or those falling closer to the middle of the distribution (Opper, 2019).

If your postsecondary success model will be used in high-stakes contexts, additional limitations apply. In general, all limitations in lower-stakes contexts become even more important in higher-stakes contexts. For instance, a lack of precision in generating indicators is unfortunate in a low-stakes context such as awarding recognition for excellence, but it could lead to more severe consequences if results are used to identify low-performing schools as a part of an accountability system. Additionally, new cautions emerge because the high-stakes use of a postsecondary success model can lead to the corruption of measures and unintended consequences. For instance, models that account for certain student subgroups (for example, English learner students) that perform lower than average can create an incentive for schools to classify more students in that subgroup in order to raise their postsecondary success indicators. Similarly, models with a narrow definition of postsecondary success could reduce the emphasis on alternative pathways to postsecondary success.

Appendix A. Kansas Postsecondary Effectiveness Model

The Kansas Postsecondary Effectiveness Model (K-PEM) is a performance measure used to assess the quality of Kansas school districts in terms of student postsecondary success and to signal the importance of preparing students to succeed after high school. The Kansas State Department of Education (KSDE) plans to incorporate data generated by the K-PEM in district report cards to allow for comparisons.

The K-PEM defines student postsecondary success as the percentage of students in a given district with the same expected high school graduation year (a student cohort) who were continuously enrolled in college for two years after their expected high school graduation year or who completed a postsecondary credential within two years of high school, including earning an approved professional certification while in high school. District postsecondary success indicators are calculated using the grade 9 cohort. The calculation includes students who dropped out or transferred into the district after grade 9 and excludes students who transferred out of the district after grade 9. This same cohort is used to calculate the covariate values described below.

Student demographic covariates

The K-PEM uses student demographic covariates to generate district postsecondary success indicators. The dataset includes one row of data per district, with all student covariates aggregated to the district level. The selection of covariates was informed by a KSDE practitioner workgroup, which was tasked with identifying factors that would accurately predict low-performing districts. KSDE leaders provided further guidance on including covariates that they regarded as being largely outside districts' control. Quantitative analyses were conducted to examine the relationships between the covariates and student postsecondary success as well as the proportion of variance in student postsecondary success accounted for by models with varying combinations of covariates. The covariates retained in the final model were student chronic absenteeism (missing 10 or more days of school), mobility (transferring schools at least once during the school year), eligibility for the national school lunch program, English learner status, and disability status.

In measuring the student demographic covariates, the K-PEM accounts for their time-varying nature. For example, students' English learner status can change during their high school years. The K-PEM uses student-level data to calculate the percentage of time during which each student was classified as having a particular demographic characteristic while enrolled in a Kansas public high school. This approach uses all relevant data available for each student in the state's longitudinal data system. Students can have 1–12 years of data from which to calculate their covariate values. For example, a student who has data available only for grades 9–12 and who was classified as an English learner student only in grade 9 would be classified as an English learner student for 25 percent of the time. Student-level data are aggregated to the district level so that covariate values represent the average percentage of time students in a district had a given designation (see appendix B for additional detail on calculating aggregated longitudinal covariates).

Student cohorts

To increase the precision of district postsecondary success indicators, the K-PEM uses data from five consecutive student cohorts. For example, postsecondary indicators for 2015 are generated using data from the 2010/11–2014/15 graduation cohorts (table A1). To calculate all variables included in the model, data across all five cohorts are pooled, resulting in a dataset in which each district has a single value for each variable that is entered into the model. This dataset is then used to generate district postsecondary success indicators.

The K-PEM is a single-level, residual model that uses a linear regression approach to generate a predicted postsecondary success score for each district. These predicted postsecondary success scores are compared with actual (observed) scores to calculate a residual score for each district (the difference between a district's actual and predicted scores). KSDE uses the residual score as an indicator of the district's postsecondary success.

A district whose actual postsecondary success score is greater than its predicted postsecondary success score is considered to be effective. The Postsecondary Effectiveness Star Award, part of the Kansans Can Star Recognition Program, recognizes districts that exceed their predicted postsecondary success score. The highest recognition goes to districts that exceed their predicted postsecondary success score and whose actual postsecondary success score is at or above 70 percent (Kansas State Department of Education, n.d.).

Table A1. Student cohorts and data included in the 2015 Kansas Postsecondary Effectiveness Model indicator

Student cohort	Dependent variable[a] (years postsecondary success is observed)	Covariates[b] (years demographic characteristic is observed)
2011	2012–13	Through 2011
2012	2013–14	Through 2012
2013	2014–15	Through 2013
2014	2015–16	Through 2014
2015	2016–17	Through 2015

Note: All covariates are calculated using all of each student's data that are available through the graduation year. The aggregate school-level covariates represent the percentage of student years all students in each district had a particular demographic designation.

a. Two-year postsecondary college enrollment or attainment of a certificate.

b. Covariates are calculated as the percentage of school years that a student is chronically absent, mobile, eligible for the national school lunch program, an English learner student, or has a disability.

Source: Authors' construction based on data from Moss (2017).

Appendix B. Technical details of the student postsecondary success model

Calculating time-varying covariates

You have two main options when calculating aggregated time-varying covariates. Both approaches operationalize a covariate as the average percentage of time students in a cohort, school, or district had a particular demographic designation (for example, average percentage of time students were classified as English learner students in a school). First, you could calculate the percentage of the high school years (grades 9–12) during which each student had a particular demographic designation. You calculate this variable using the number of years each student was in the given high school or district as the denominator and the number of those years during which each student had a given designation as the numerator. You then aggregate these data to the cohort, school, or district level. In this manner, all students are weighted equally regardless of how long they spent in a given school or district.

Alternatively, you could sum the total number of years all students were in each cohort, school, or district during grades 9–12. You then sum the number of years all students had a particular designation during that time. To calculate the cohort, school, or district covariate value, you divide the summed years with that designation by the total numbers of years students were in each cohort, school, or district:

Summed years with demographic designation / summed years in cohort, school, or district

This second approach can provide a more precise estimate for the cohort-, school-, or district-level covariate by essentially weighting the calculation by the number of years of available data for each student. For example, a student who spent four years in a high school would be weighted four times as much as a student who spent only one year.

Model specifications and equations

Single-level models

Single-level models use data from a single cohort of students or data from multiple cohorts that have been pooled. Each school has a single data point for each variable included in the model. With this approach, a single-level linear regression represents the most straightforward way to model school postsecondary success. The regression will generate a single postsecondary success indicator for each school.

Multilevel models

Multilevel models use hierarchical linear modeling in which each student cohort is associated with its school, and the correlations between each cohort's data are calculated and accounted for. That method accounts more appropriately for the variability that occurs within a given cluster (such as a school) and therefore produces a more precise estimate for a given school and better accounts for uncertainty (Raudenbush & Bryk, 2011). For this approach, in which cohorts are nested in schools, dummy-coded variables representing the cohorts are entered into level 1 of the statistical model, while the school is represented in level 2. This approach includes cohort (time) as a fixed-effect in the statistical model. This model will generate a separate postsecondary success indicator for each school's cohorts included in the model.

Fixed-effects approach

The fixed-effects approach includes a dummy variable for each school. A school's postsecondary success indicator is the regression coefficient associated with that school's dummy variable. The school's regression coefficient is the estimated relationship between that school and student postsecondary success, accounting for all covariates in the model. In other words the school regression coefficients represent the estimated influence of each school on student postsecondary success, holding all the covariates constant (statistically equating all schools based on their covariate values). A regression coefficient of zero indicates that the school has no association with student postsecondary success after all covariates are accounted for, a positive coefficient indicates a positive association, and a negative coefficient indicates a negative association. The larger the coefficient, the larger the school's estimated influence on student postsecondary success.

Residual approach

The residual approach uses a school's covariate values to predict the student postsecondary success of each school. The model calculates the overall influence of each covariate on students' actual postsecondary success, across all schools. The model then uses each school's covariate values to predict the school's postsecondary success score—an estimate of how well a school would be expected to perform given its student characteristics. The predicted score is then subtracted from the actual or observed postsecondary success score. This residual score is the value of that school's postsecondary success indicator. A positive residual suggests that the school performed better than expected given its student demographic or community characteristics. A negative residual suggests that the school performed worse than expected.

Example model equations

Four simplified example equations are provided below for the four model specifications described in the tool. The equations represent the combination of options you have for using a single-level or a multilevel model and a fixed-effects or a residual approach.

Single-level fixed-effects model

$$Y_i = \beta_0 + \beta_{1i}(\textit{covariates}) + \beta_2(\textit{school/district dummy variables}) + e_i$$

where Y_i is the observed score of school/district *i*, β_0 is the intercept common to all schools/districts, β_{1i} is a matrix of regression coefficients for all covariates included in the model for school/district *i*, β_2 is a matrix of regression coefficients for all school/district dummy variables, and e_i is the error term for school/district *i*

Single-level residual model

$$Y_i = \beta_0 + \beta_{1i}(\textit{covariates}) + e_i$$

where Y_i is the observed score of school/district *i*, β_0 is the intercept common to all schools/districts, β_{1i} is a matrix of regression coefficients for all covariates included in the model for school/district *i*, and e_i is the error term for school/district *i*

Multilevel fixed-effects model

$$\text{Level 1 (cohort): } Y_{ci} = \pi_{0i} + \pi_{1i}(\textit{covariates}) + \pi_{2i}(\textit{cohort dummy variables}) + e_{ci}$$

$$\text{Level 2 (school/district): } \pi_{0i} = \beta_{00} + \beta_{01}(\textit{school/district dummy variables})_i + r_{0i}$$

$$\pi_{2i} = \beta_{01}(\textit{school/district dummy variables})$$

where Y_{ci} is the observed score of cohort *c* at school/district *i*, π_{0i} is the intercept common to all student cohorts at school/district *i*, π_{1i} is a matrix of regression coefficients included for each cohort at school/district *i*, π_{2i} is a matrix of regression coefficients for all student cohort dummy variables at school/district *i*, e_{ci} is the error term for cohort *c* at school/district *i*, β_{00} is the intercept common to all schools/district, β_{01} is a matrix of regression coefficients for all school/district dummy variables, and r_{0i} is the school error term for school/district *i*

Multilevel residual model

$$\text{Level 1 (cohort): } Y_{ci} = \pi_{0i} + \pi_{01}(\textit{cohort dummy variables})_i + \pi_{1i}(\textit{covariates}) + e_{ci}$$

$$\text{Level 2 (school/district): } \pi_{0i} = \beta_{00} + r_{0i}$$

where Y_{ci} is the observed score of cohort *c* at school/district *i*, π_{0i} is the intercept common to all student cohorts at school/district *i*, π_{01} is a matrix of regression coefficients for all student cohort dummy variables at school/district *i*, π_{1i} is a matrix of regression coefficients included for each cohort at school/district *i*, e_{ci} is the error term for cohort *c* at school/district *i*, β_{00} is the intercept common to all schools/districts, and r_{0i} is the school/district random effect.

References

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college*. U.S. Department of Education. https://eric.ed.gov/?id=ED490195.

Berk, R. A. (1976). Determination of optimal cutting scores in criterion–Referenced measurement. *Journal of Experimental Education, 45*(2), 4–9. https://doi.org/10.1080/00220973.1976.11011567.

Caro, D. H., Cortina, K. S., & Eccles, J. S. (2015). Socioeconomic background, education, and labor force outcomes: Evidence from a regional US sample. *British Journal of Sociology of Education, 36*(6), 934–957. https://eric.ed.gov/?id=EJ1073727.

Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Sage.

Conley, D. T. (2014). New conceptions of college and career ready: A profile approach to admission. *Journal of College Admission, 223*, 12–23. Retrieved July 24, 2020, from https://www.erblearn.org/news/new-conceptions-college-and-career-ready-profile-approach-admission.

Deutsch, J., Johnson, M., & Gill, B. (2020). *The promotion power impacts of Louisiana high schools.* Mathematica. https://www.mathematica.org/our-publications-and-findings/publications/the-promotion-power-impacts-of-louisiana-high-schools.

Gheen, M., Smerdon, B., Hein, V., & Lebow, M. (2012). *Outcomes and measures for college and career success: How do we know when high school graduates meet expectations*. National High School Center at the American Institutes for Research. Retrieved July 24, 2020, from https://ccrscenter.org/products-resources/outcomes-and-measures-college-and-career-success-how-do-we-know-when-high-school.

Hein, V., Smerdon, B., & Sambolt, M. (2013). *Predictors of postsecondary success*. College and Career Readiness and Success Center at American Institutes for Research. https://eric.ed.gov/?id=ED555671

Kansas State Department of Education. (n.d.). *Postsecondary effectiveness rate*. Retrieved July 24, 2020, from https://www.ksde.org/Agency/Fiscal-and-Administrative-Services/Communications-and-Recognition-Programs/Vision-Kansans-Can/Kansans-Can-Star-Recognition/Postsecondary-Effectiveness-Rate

Moss, T. (2017). *Technical questions & current specifications: Using predicted district effectiveness rates as a performance benchmark*. Kansas State Department of Education. https://eric.ed.gov/?id=ED589795

National Student Clearinghouse. (2016). *National college progression rates for high schools participating in the National Student Clearinghouse StudentTracker service*. Retrieved July 24, 2020, from https://nscresearchcenter.org/hsbenchmarks2016/

Newton, X. A., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Educational Policy Analysis Archives, 18*(23), 1–27. https://eric.ed.gov/?id=EJ913473.

Opper, I. M. (2019). *Value-added modeling 101: Using student test scores to help measure teaching effectiveness.* RAND Corporation. https://www.rand.org/pubs/research_reports/RR4312z1.html.

Raudenbush, S. W., & Bryk, A. S. (2011). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage.

Rose, R. A., Henry, G. T., & Lauen, D. L. (2012). *Comparing value-added models for estimating individual teacher effects on a statewide basis: Simulations and empirical analyses*. Consortium for Educational Research and Evaluation–North Carolina. https://publicpolicy.unc.edu/files/2015/07/Comparing-value-added-models-for-estimating-individual-teacher-effects-on-a-statewide-basis-1.pdf

REL 2021–074

March 2021

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-17-C-0005 by the Regional Educational Laboratory Central administered by Marzano Research. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Brodersen, R. M., Gagnon, D., Liu, J., & Moss, T. (2021). *Steps to develop a model to estimate school- and district-level postsecondary success* (REL 2020–074). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Central. http://ies.ed.gov/ncee/edlabs.

This resource is available on the Regional Educational Laboratory website at http://ies.ed.gov/ncee/edlabs.