IES Institute of Education Sciences

# An Approach to Using Student and Teacher Data to Understand and Predict Teacher Shortages

Regional Educational Laboratory Central

At Marzano Research

REL 2021–052
U.S. DEPARTMENT OF EDUCATION

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

# An Approach to Using Student and Teacher Data to Understand and Predict Teacher Shortages

*Robert Reichardt, Mary Klute, Joshua Stewart, and Stephen Meyer*

December 2020

**This resource describes an approach to identifying patterns of teacher shortages that was developed collaboratively by the Missouri Department of Elementary and Secondary Education and Regional Educational Laboratory Central. The approach uses widely available software and data and can be adopted or adapted by other education agencies that wish to understand and predict teacher shortages.**

# Contents

# Why this resource?

Addressing teacher shortages, especially in particular subject areas, grade levels, and geographic locations, has been a persistent concern among leaders in schools, districts, state education agencies, and the federal government (Borman & Dowling, 2008; Malkus et al., 2015). Shortages occur when the number of teachers available to work given current wages and working conditions (or the supply of teachers) is smaller than the number of available teaching positions (the demand for teachers; Boe & Gilford, 1992).

Recent studies have suggested a large decrease over the past decade in enrollment in teacher preparation programs, an important source of teacher supply (U.S. Department of Education, 2015), and projected a substantial national teacher shortage over the next decade (Sutcher et al., 2016). Other recent national and state studies have indicated that teacher shortages differ substantially across subject areas and geographic locations (Meyer et al., 2019; Pennington McVey & Trinidad, 2019).

Teacher shortages can limit students' access to high-quality teachers, which is associated with reduced student success. Shortages can lead to vacancies being filled by teachers who are not appropriately certified, which can be problematic because those educators might not have the necessary skills to teach courses with more complex content, such as Algebra I (Stewart et al., 2019). In Missouri 16.3 percent of teachers in high-poverty schools statewide were not appropriately certified compared with 5.7 percent of teachers in schools with fewer students from low-income households (Missouri Department of Elementary and Secondary Education, 2015).

Teacher shortage data can be used in developing strategies to mitigate inequities in students' access to high-quality teachers. However, state reports of teacher shortages vary widely in the complexity of their data sources and analysis (Education Commission of the States, 2019). Some state education agencies use a single data source such as survey data from schools and districts; others use multiple sources, including administrative data and information about teacher supply from teacher preparation programs (Aldeman, 2018; Lindsay et al., 2009). Some agencies, often with assistance from external researchers, go beyond federal reporting requirements and conduct research that informs state-specific plans and responses to teacher shortages or surpluses. For example, analyses of state administrative data have described the current teacher workforce, including teachers' demographics, the locations in which they work, the subjects they teach, and their qualifications (Aldeman, 2018; Folsom et al., 2014; Lindsay et al., 2009; Lindsay et al., 2016; Reichardt, 2003; White & Fong, 2008). Fewer analyses have used state administrative data to make predictions in order to help plan responses to future teacher demand and shortages (Aldeman, 2018; Berg-Jacobson & Levin, 2015; Levin et al., 2015; Lindsay et al., 2009; Reichardt, 2003).

State education agencies can also use predictions of teacher shortages to address inequities in students' access to high-quality teachers. The importance of teacher effectiveness is well supported by studies on how teacher ability contributes to student achievement gains. All else being equal, students taught by more qualified teachers experience greater achievement gains than do students taught by less qualified teachers (Aaronson et al., 2007; Chetty et al., 2014; Konstantopoulos & Chung, 2011; Nye et al., 2004). Some research points to inequities in access to high-quality teachers, finding that students who attend high-poverty, high–racial/ethnic minority schools are, to varying extents, more likely to be taught by less effective teachers with fewer qualifications and less experience (Clotfelter et al., 2010; Glazerman & Max, 2011; Isenberg et al., 2013; Isenberg et al., 2016; Sass et al., 2012).

Like many other state education agencies, the Missouri Department of Elementary and Secondary Education wants to analyze teacher shortages to ensure equitable access to high-quality teachers but has limited resources to do so without external support. The department and the Regional Educational Laboratory Central collaborated to identify a relevant, understandable, and rigorous approach to describing historical and current demand for teachers and predicting shortage areas in Missouri, including vacancy trends by grade level and certification area. The result was the teacher predictor model, which was designed for department staff to use on their own with Microsoft Excel and existing administrative data. The information from the model can guide decisions throughout the teacher pipeline by teacher preparation programs, school districts, and state administrators and policymakers. The information can also be used to make potential teachers aware of future opportunities and to support planning and teacher preparation programs in areas with predicted shortages and to help districts meet federal reporting requirements.

An important lesson from this work is that regional leaders must be able to understand the basic structure of the model to have confidence in using the results. The model cannot be a "black box" if its results are to be used (Armstrong et al., 2011).

This resource documents how the teacher predictor model was developed and implemented in Missouri and demonstrates how it makes predictions so that other education agencies can adopt or adapt the approach to understand and predict teacher shortages in their own context. The resource describes the two main processes in the model: predicting student enrollment and predicting the teacher workforce, assuming current trends continue. Users can choose from several methods for making predictions, including using the most recent year of data, using averages of previous years of data, and using a regression-based method.

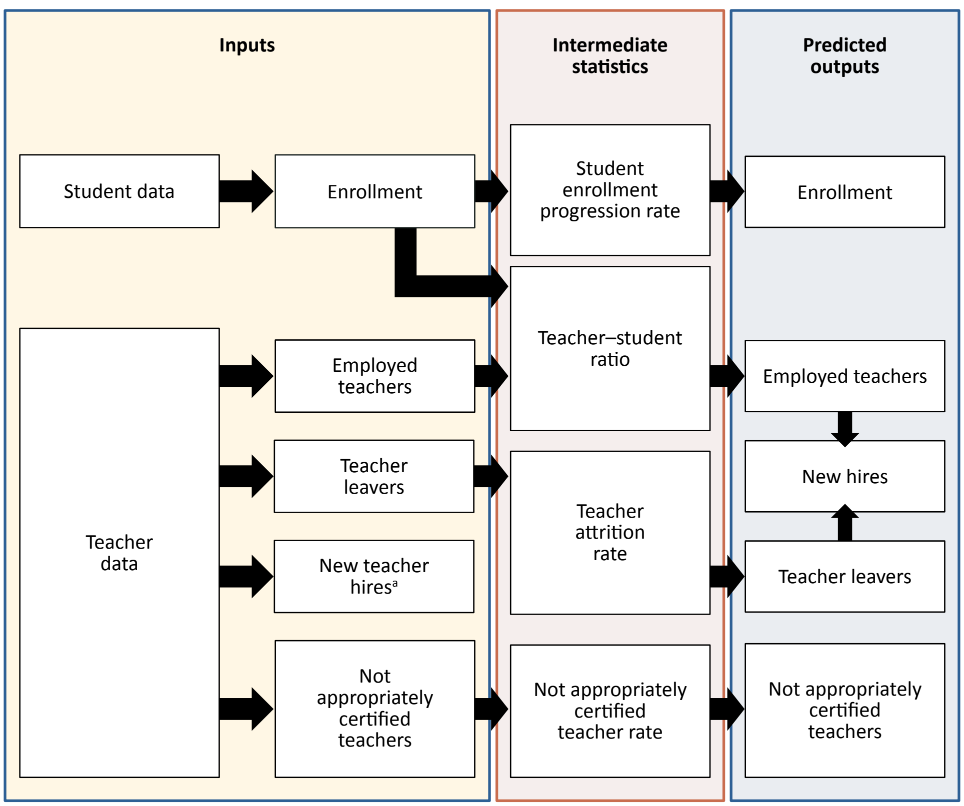
# What is the teacher predictor model?

The teacher predictor model uses student enrollment and teacher assignment data from recent academic years to make predictions for future academic years. It has three main components: inputs, intermediate statistics, and predicted outputs (figure 1). Inputs consist of historical and current student enrollment and teacher data, including counts of all employed teachers, teachers who leave their districts (teacher leavers), new hires, and teachers who are not appropriately certified. The inputs are used to calculate four intermediate statistics: student enrollment progression rate, teacher–student ratio, teacher attrition rate, and not appropriately certified teacher rate. The intermediate statistics are used to calculate the model outputs: predicted student enrollment and predicted teacher workforce, including the numbers of employed teachers, teacher leavers, new hires, and not appropriately certified teachers. In the Missouri model the number of not appropriately certified teachers was used as a shortage indicator (see box 1 for definitions of key terms used in this resource).

The model predicts the counts of employed teachers, teacher leavers, new hires, and not appropriately certified teachers for each region of a state. Regions are used to reflect the local nature of labor markets and shortages (Boyd et al., 2005). In the Missouri model Kansas City and St. Louis City, the two largest urban districts in Missouri, were separated into their own regions because their labor markets differ from those of neighboring suburban districts.

Several methods can be used to calculate the intermediate statistics. To inform decisions about which method or methods to use, this resource describes a validation process that uses historical data to make predictions and then assesses their accuracy by comparing them with actual data.

**Figure 1. Representation of the teacher predictor model**




a. Teacher data on new teacher hires are not used in the prediction calculations. These data are used for model validation (that is, to compare the predicted number of new teacher hires with actual data).

Source: Authors' creation.

## Box 1. Key terms

The key terms used in this resource refer to the teacher predictor model developed for the Missouri Department of Elementary and Secondary Education. Although many terms and definitions are generalizable for models in other states, they may vary according to state-specific factors such as available data, policies, and priorities.

**Average percentage error.** A metric used to assess the validity of the model's intermediate statistics on the basis of whether predictions are biased in a certain direction. It is the sum of the differences between predicted and actual values, divided by the sum of actual values. The differences are the errors, which are calculated for each grade level and subject area by region and statewide.

**Employed teacher.** A staff member who is assigned the professional activities of instructing students in prekindergarten–grade 12 in self-contained classes or courses in a public school.

**Grade level.** The grade band in which a teacher taught for the majority of an academic year. In the Missouri model, grade level was designated as elementary school (prekindergarten–grade 5), middle school (grades 6–8), or high school (grades 9–12).

**Mean absolute percentage error.** A metric used to assess the validity of the model's intermediate statistics on the basis of the size of the errors in the predictions. It is the sum of the absolute value of the differences between predicted and actual values, divided by the sum of actual values. The differences are the errors, which are calculated at the subject area and grade levels for each region and statewide.

**Intermediate statistics.** The four annual rates and ratios that are the basis of the model's predictions: student enrollment progression rate, teacher–student ratio, teacher attrition rate, and not appropriately certified teacher rate.

**New hire.** A teacher who did not teach in the same district during the prior academic year.

**Not appropriately certified teacher.** A designation for an individual who meets the minimum qualification to teach but lacks the appropriate subject-area credits to teach in his or her assigned grade level and subject area, based on state certification rules. This definition varies according to state certification rules.

**Not appropriately certified teacher rate.** The number of teachers who are not appropriately certified divided by the total number of teachers. This rate uses a Missouri-specific determination of whether a teacher has the appropriate subject-area credits and certification to teach a given grade level and subject area. This definition varies by state. In the Missouri model not appropriately certified teacher rates were calculated annually for each region, grade level, and subject area.

**Primary assignment.** The subject area, grade level, and district in which a teacher spends the majority of time in an academic year.

**Region.** The geographic region of the state in which a teacher has a primary assignment.

**Student enrollment data.** The count of students by year, grade level, and district.

**Student enrollment progression rate.** The number of students enrolled in a grade divided by the number of students enrolled in the prior grade in the previous year. In the Missouri model student enrollment progression rates are calculated annually for each county, grade level, and subject area.

**Subject area.** The subject area in which a teacher has a primary assignment.

**Teacher attrition rate.** The number of teacher leavers divided by the total number of teachers. In the example presented in this resource, rates are calculated annually for each region, grade level, and subject area.

**Teacher leaver.** A teacher with a primary assignment to a district in an academic year who does not have a primary assignment to that district in the subsequent year.

**Teacher shortage.** When the demand for teachers exceeds the supply. Missouri's model used predicted rates of not appropriately certified teachers in particular regions, grade levels, and subject areas as an indicator of teacher shortages (though the rate can also be used to predict the number of not appropriately certified teachers).

**Teacher–student ratio.** The number of employed teachers divided by the number of enrolled students. In the Missouri model teacher–student ratios were calculated annually for each region, grade level, and subject area.

The teacher predictor model uses data that are commonly available in state administrative data systems. The Missouri model was applied using data from the Missouri Student Information System, which contains historical and current individual-level data on students and educators in Missouri public schools. To demonstrate how the model is designed, this resource uses two types of data for academic years 2011 through 2018:

- Student enrollment by grade and district.
- Teacher data, including not appropriately certified designation by the Missouri Department of Elementary and Secondary Education, grades taught, subject areas taught, and district assignments.

# Model assumptions

The teacher predictor model relies on three main assumptions. If these assumptions are not tenable, the model may not provide accurate predictions.

First, the model assumes that input data are accurate. Because the accuracy of predictions hinges on data quality, users should consider the quality of available data when developing the model and evaluating its results. The state education agency data used in teacher prediction models are typically provided by districts and may include errors, reducing the accuracy of predictions. Although state education agencies can validate the demographic and certification data that districts provide, teacher assignment data are often more difficult to validate.

Second, the model assumes that historical patterns will continue into the future. The model also assumes that future patterns are linear and unchanging over the prediction period, which was four years in the Missouri model.[1] For example, the model assumes that rates of not appropriately certified teachers are stable over time. However, the model allows for nonlinear changes to enrollment that can impact predictions. The model may be less accurate when there are large changes in demographic trends, the economy, fiscal policy, salaries, benefits, or education policy; users might wish to examine predictions more closely under such circumstances. The accuracy of predictions may decrease the further out they are in the future.

Finally, the model assumes that the positions occupied by teachers who are not appropriately certified would be filled by appropriately certified teachers if they were available. This assumption is common in the literature—for example, a study in Minnesota on teachers for whom districts had to apply for "special permissions" to fill vacant positions (Lindsay et al., 2016) and a study in Oklahoma on teachers with emergency certifications (Berg-Jacobsen & Levin, 2015). In Missouri, state administrative data include information on teachers who are not appropriately certified, and the Missouri model used the not appropriately certified rate as an indicator of teacher shortages. An advantage of this indicator is that it is verified through the state's Core Data System and is therefore reliable. A disadvantage is that it may underestimate shortages. Because the model's shortage predictions are based on positions filled with not appropriately certified teachers, they do not include positions for which no teachers are available to teach regardless of qualifications. In other words, this indicator does not identify shortages for courses that are not offered at all because districts cannot find teachers to teach them.

1. The Missouri Department of Elementary and Secondary Education chose a four-year prediction period in order to provide enough time for an educator preparation program to influence new potential teachers entering college to select an education major associated with a shortage area.

# How does the teacher predictor model make predictions?

The key output from the teacher predictor model is the predicted number of not appropriately certified teachers, which is used as the indicator of teacher shortages. This section starts by describing how the model predicts student enrollment. It then discusses how the model predicts the teacher workforce, including the predicted number of not appropriately certified teachers, new hires, and teacher leavers.

Users must choose the method to calculate the intermediate statistics for the predications of student enrollment and teacher workforce. The Missouri Department of Elementary and Secondary Education considered four methods to calculate the intermediate statistics: most recent year, two-year average, four-year average, and a linear trend line function. After describing how the model works conceptually, this section discusses these methods as well as the process for validating the selected method. Step-by-step instructions for making the calculations in the model are in appendix A.

## Predicting student enrollment

The first process in the teacher predictor model uses the student data input (historical and current enrollment) to calculate one intermediate statistic: the student enrollment progression rate. The annual progression rate is then used to predict student enrollment (see figure 1).

### Inputs

The inputs for predicting student enrollment include data on enrollment in prior and current years. Counts of enrolled students are disaggregated by county, grade, and year. The Missouri Department of Elementary and Secondary Education used a spreadsheet to organize these data.

### Intermediate statistics

The process for predicting student enrollment involves one intermediate statistic: the student enrollment progression rate. The model predicts student enrollment by using the cohort progression method, which was found to be the most accurate after predictions were compared with actual enrollment data (Lindsay et al., 2016; Minnesota Department of Education, 2015).

The cohort progression method involves applying an anticipated progression rate, calculated using historical data, to current cohorts of students. The numbers of students by grade for each academic year are summed by county and used to determine progression rates of students from one grade to the next (for example, from grade 1 to grade 2). Counties are used because they allow for aggregation of smaller districts to minimize very small cell sizes seen in some rural districts while still capturing local variation in enrollment patterns. These progression rates are applied to the most recent enrollment data to predict student enrollment in future years.

A mathematical expression of the approach is shown below. Student enrollment data are used to estimate the number of students who progress from one grade to the next. County-level progression rates based on enrollment counts of students from year to year are used rather than the progression of individual students. Progression rates are calculated by dividing enrollment for the current year in each grade by enrollment for the previous year in each prior grade. In year $y$ and grade $k$ the annual progression rate ($APR$) is the enrollment ($E$) in year $y + 1$ and grade $k + 1$, divided by the enrollment in year $y$ and grade $k$:

$$\text{APR}_{y+1,k+1} = E_{y+1,k+1} / E_{yk}.$$

The model predicts the number of students in grades 1–12 in the next year by multiplying enrollment in the prior year and grade by these progression rates.

Figure 2 illustrates the cohort progression method. The arrows represent the progression rates, describing the proportion of students who advance from one grade to the next. For example, the 2016 progression rate for kindergarten students is the number of grade 1 students in 2017 divided by the number of kindergarten students in 2016. This calculation yields the number of students who advanced from kindergarten to grade 1 from a given year to the next.

**Figure 2. Illustration of the cohort progression method used to predict grade 1–12 student enrollment from 2016 to 2017**

| | | | | |
|---|---|---|---|---|
| **2017** | **K** | **1** | **2** | **3** |
| **2016** | **K** | **1** | **2** | **3** |

Source: Authors' creation.

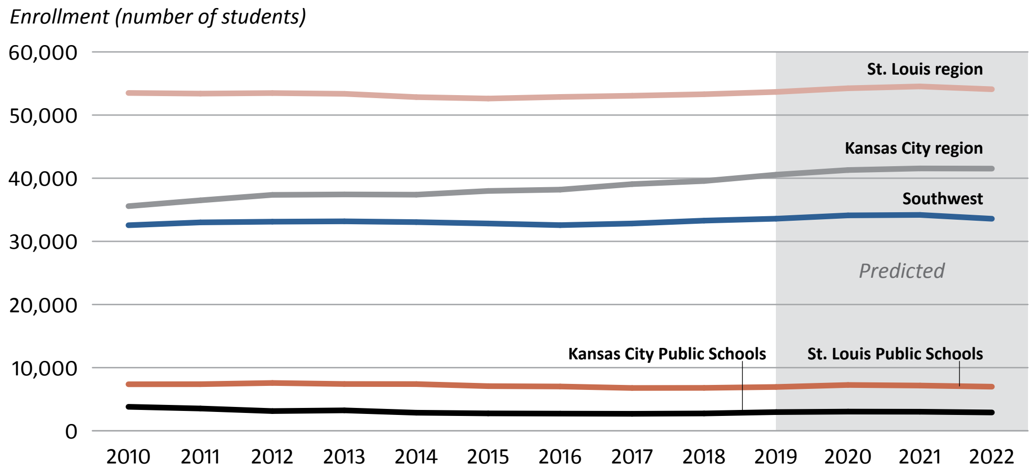
Predicting enrollment for prekindergarten and kindergarten is less straightforward. Prekindergarten is the first year in the model, which means progression rates that rely on a prior year of data cannot be used. Available options include historical enrollment rates (Pettibone & Bushan, 1990; Sweeney & Middleton, 2005), birth-to-kindergarten ratios (Lindsay et al., 2016), and fertility rates (Levin et al., 2015). Options that have a substantial time lag, such as the birth-to-kindergarten ratio, are less useful when there have been recent changes in enrollment patterns. Because prekindergarten enrollment has recently increased in Missouri, the Missouri Department of Elementary and Secondary Education chose to use historical enrollment counts to predict the future number of prekindergarten and kindergarten students.

## Predicted outputs

To calculate predicted student enrollment, the annual student enrollment progression rate is multiplied by observed student enrollment. Figure 3 provides an example, showing historical and predicted public middle school enrollment for selected Missouri regions. Enrollment in several regions, particularly the Kansas City region, is predicted to continue to increase while enrollment in Kansas City Public Schools and St. Louis Public Schools is predicted to remain flat.

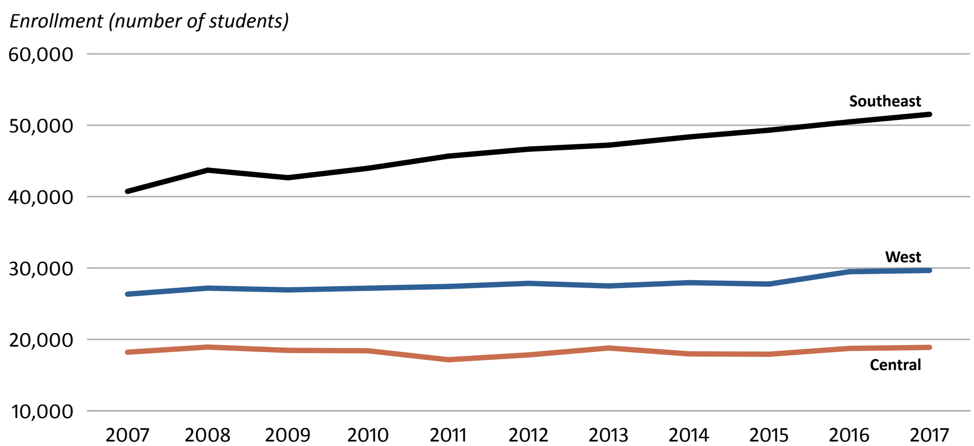
The trends depicted in figure 3 are relatively flat. In contrast, figure 4 provides an example of more dynamic trends in growth in student enrollment for selected South Dakota regions. The Southeast region has had rapid growth, whereas growth has been slower in the West and Central regions.

**Figure 3. Historical and predicted middle school enrollment for selected Missouri regions, 2010–22**




Source: Authors' construction using data from the Missouri Department of Elementary and Secondary Education.

**Figure 4. Historical enrollment for selected South Dakota regions, 2007–17**




Source: Authors' construction using data from the U.S. Department of Education's Common Core of Data.

# Predicting the future teacher workforce

The second process of the teacher predictor model uses the teacher data inputs to calculate three additional intermediate statistics: the teacher–student ratio, the teacher attrition rate, and the not appropriately certified teacher rate. These intermediate statistics are then used to predict the future numbers of employed teachers, teacher leavers, and not appropriately certified teachers and the need for new hires (see figure 1).

## Inputs

The four model inputs for predicting the teacher workforce are counts of employed teachers, teacher leavers, new hires, and not appropriately certified teachers. These data included information about where teachers were employed and whether teachers were appropriately certified each academic year. The Missouri Department of Elementary and Secondary Education summed the individual records for each combination of region, grade level, and subject area (for example, the count of elementary music teachers in the Central region).

Table 1 provides an example of those totals, showing the number of teachers with each subject area assignment in the Central region of Missouri. The largest single subject area is elementary (1,679), followed by special education (895) and English language arts, journalism, speech, and dramatics (716).

The counts of employed teachers for each combination of region, grade level, and subject area offer descriptive information about the teacher workforce, such as the numbers of employed high school teachers in selected subject areas over several years. Figure 5 provides an example of information on recent trends in staffing that stakeholders could gain from these data. Because the workforces are large, changes were expected to be small over the short time frames shown. Subject areas were selected to show diversity in counts, and data are depicted for every other year because counts change slowly. For example, from 2012 to 2018 the number of music teachers increased slightly, and the number of special education and business teachers decreased slightly.

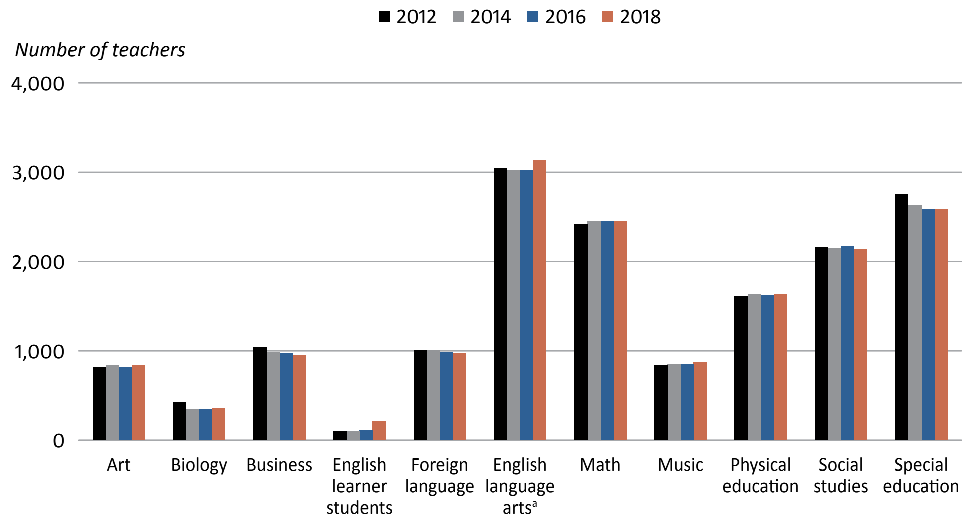
**Table 1. Number of employed teachers by grade level and subject area for the Central region of Missouri, 2015**

| | Grade level | | | |
|---|---|---|---|---|
| Subject area | Elementary | Middle | High | Total |
| Agriculture | † | 6 | 57 | 63 |
| Art | 70 | 43 | 81 | 194 |
| Biology | † | † | 29 | 29 |
| Business | † | 23 | 88 | 111 |
| Chemistry | † | † | 20 | 20 |
| Earth science | † | 6 | 4 | 10 |
| Elementary | 1,679 | † | † | 1,679 |
| English language arts, journalism, speech, and dramatics | 197 | 224 | 295 | 716 |
| English learner students | 23 | 11 | 11 | 45 |
| Family consumer science and human service education | † | 23 | 50 | 73 |
| Foreign language | † | 24 | 81 | 105 |
| Health occupations | † | † | 2 | 2 |
| Industrial technology | † | 9 | 23 | 32 |
| Marketing and cooperative education | † | † | 7 | 7 |
| Math | 17 | 163 | 231 | 411 |
| Music | 69 | 64 | 80 | 213 |
| Physical education and health | 81 | 103 | 166 | 350 |
| Physics | † | † | 6 | 6 |
| Science | 1 | 139 | 164 | 304 |
| Skilled technical science | † | † | 4 | 4 |
| Social studies | 2 | 154 | 210 | 366 |
| Special education | 488 | 172 | 235 | 895 |

† indicates combinations of subject areas and grade levels without employed teachers.

Source: Authors' construction using data from the Missouri Department of Elementary and Secondary Education.

**Figure 5. Number of employed high school teachers by year and subject area in Missouri, 2012–18**




a. Includes English language arts, journalism, speech, and dramatics.

Source: Authors' construction using data from the Missouri Department of Elementary and Secondary Education.

## Intermediate statistics

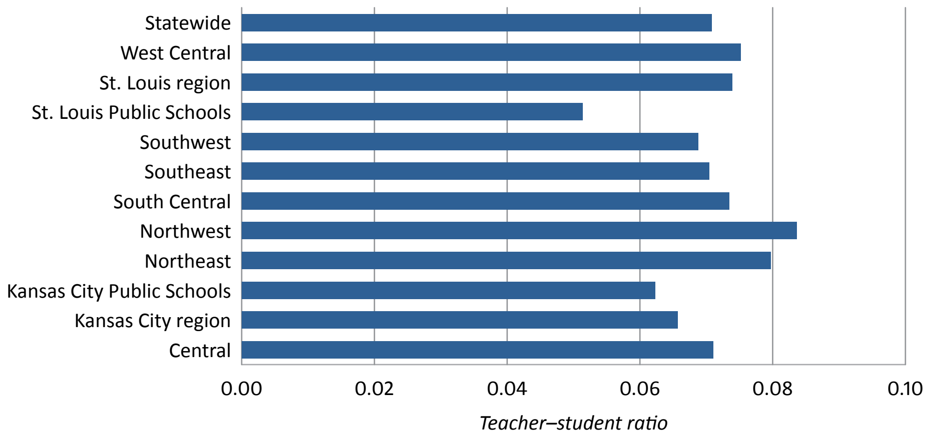
The process for predicting the future teacher workforce involves three intermediate statistics: teacher–student ratio, teacher attrition rate, and not appropriately certified teacher rate (see figure 1). This section describes different options for calculating the intermediate statistics. The Validating intermediate statistics section that follows describes how the calculation method is selected.

The teacher–student ratio combines teacher data with student enrollment counts by grade level and region to predict the number of employed teachers by region, grade level, and subject area. It is calculated by dividing the number of employed teachers by the number of enrolled students. In the Missouri model ratios were calculated for each calendar year from 2011 through 2018.

Although the teacher–student ratio was created as a model calculation and is not regularly analyzed by stakeholders, it is a metric for measuring resource allocation: it shows the proportion of teaching staff assigned to each student. It is the inverse of the student–teacher ratio: a teacher–student ratio of .05 is the same as a student–teacher ratio of 20:1. Figure 6 provides an example, showing teacher–student ratios by Missouri region and how more teaching resources were assigned to elementary schools in the Northwest region than in St. Louis City Schools.

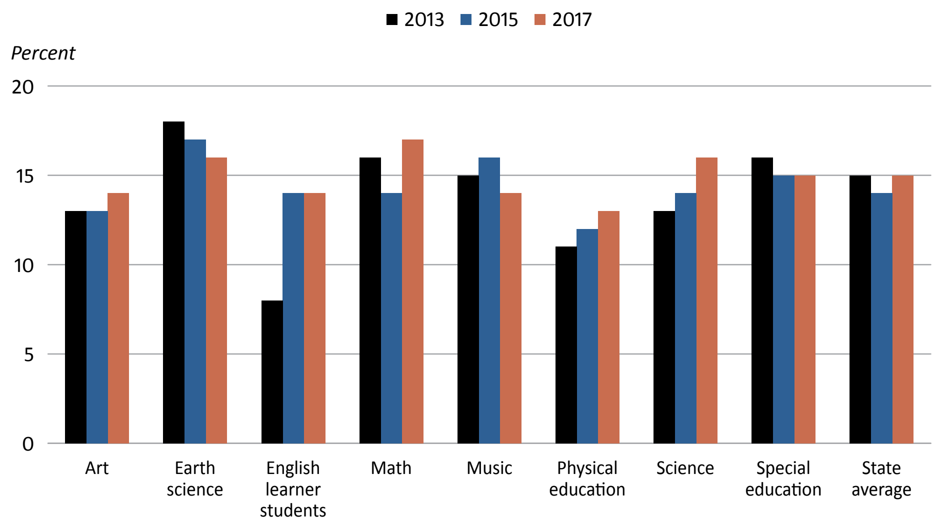
Teacher attrition rates are calculated for each academic year by dividing the number of teachers who left their districts by the number of employed teachers. Figure 7 provides an example, showing middle school attrition rates in Missouri by subject area for a five-year period. In the Missouri model attrition was defined as leaving a district because district leaders were a key audience for the model results. Because calculating the attrition rate requires two subsequent years of data, it was not calculated for the final year. The attrition rate for teachers of English learner students increased from 2013 to 2015. Some rates, such as those for teachers of earth science, were consistently higher than those for teachers of other subject areas, whereas other rates, such as those for teachers of art and teachers of physical education, were consistently lower.

**Figure 6. Elementary school teacher–student ratios by Missouri region, 2017**




Source: Authors' construction using data from the Missouri Department of Elementary and Secondary Education.

**Figure 7. Middle school teacher attrition rates by year and subject area in Missouri, 2013–17**



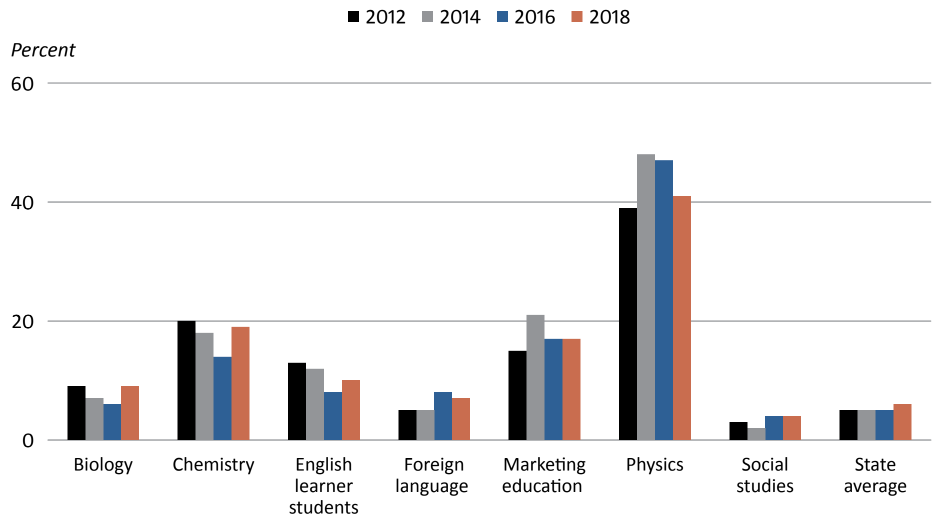

Source: Authors' construction from data from the Missouri Department of Elementary and Secondary Education.

Not appropriately certified teacher rates are calculated for each academic year by dividing the number of teachers who are not appropriately certified by the total number of employed teachers for each combination of region, grade level, and subject area. Figure 8 provides an example, showing not appropriately certified high school teacher rates for Missouri by subject area over time. Some subject areas, such as chemistry and physics, had higher rates of not appropriately certified teachers than did other subject areas, such as biology. This result suggests that challenges in recruiting qualified teachers differ by science subject.

Other rates can contextualize shortage predictions based on the not appropriately certified teacher rate. For example, high teacher attrition rates have been connected with increased teacher shortages and negative impacts on students (Ingersoll, 2001; Ronfeldt et al., 2013).

Figure 9 provides an example, plotting the teacher attrition rate alongside the not appropriately certified teacher rate for elementary school teachers in the subject areas with the most teachers in Missouri. Examining the two rates for each subject can provide context. For example, if both rates are well above the state average, it can be

**Figure 8. Not appropriately certified high school teacher rates in Missouri, by year and subject area, 2012–18**

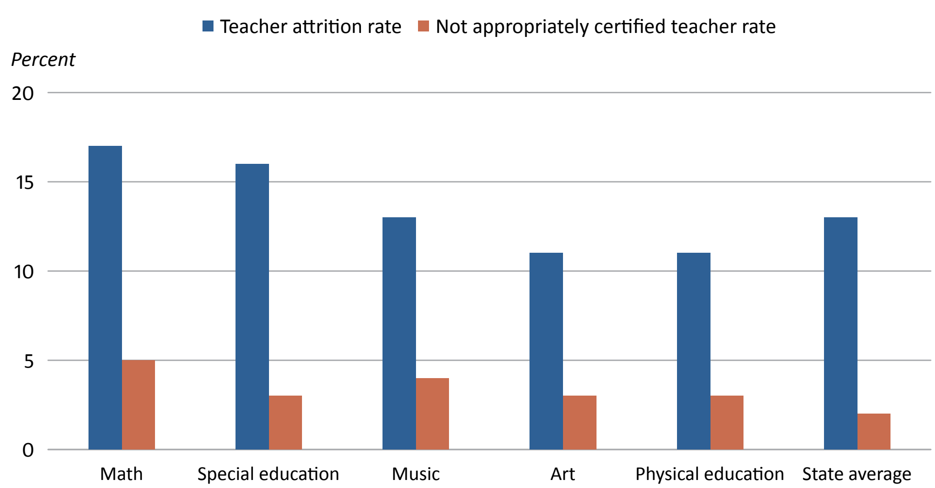



Source: Authors' construction using data from the Missouri Department of Elementary and Secondary Education.

interpreted as supporting evidence for the shortage prediction. If the two rates diverge (for example, one is well above the state average and the other is not), it could indicate that more evidence is needed on the extent of shortages in a certain subject or grade. If both rates are low, it might be evidence that a shortage does not exist.

Figure 9 shows that the not appropriately certified teacher rate is higher in math than in other subjects, which suggests a shortage of math teachers. That conclusion is reinforced by the high attrition rate in math. The figure also shows that the not appropriately certified teacher rate is lower for special education than for math. As previously noted, the not appropriately certified teacher rate does not fully capture shortages when teachers are not available to fill positions. The high attrition rate for special education teachers suggests that the not appropriately certified teacher rate does not fully capture the challenges of filling special education positions. This conclusion is supported by information from Missouri stakeholders and others about persistent challenges with hiring qualified replacement teachers.

**Figure 9. Using teacher attrition rates to contextualize not appropriately certified teacher rates by subject area among Missouri elementary school teachers, 2018**




Source: Authors' construction using data from the Missouri Department of Elementary and Secondary Education.

## Predicted outputs

The intermediate statistics are used to produce the model outputs: the predicted numbers of employed teachers, teacher leavers, new hires, and not appropriately certified teachers in future years. To predict the number of employed teachers, the teacher–student ratios for each region, grade level, and subject area are multiplied by predicted student enrollment for the same region, grade level, and subject area. The number of years for which predictions are made must balance the needs of decisionmakers with the fact that the accuracy of predictions declines over time. The Key considerations section later in this resource provides guidance on choosing the time horizon for predictions.

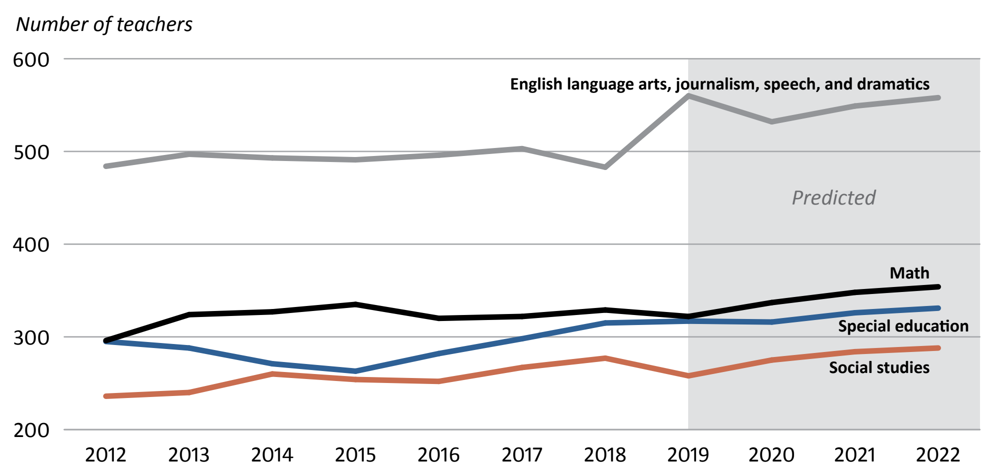
Figure 10 provides an example, showing historical and predicted numbers of employed middle school teachers in selected subject areas in the Kansas City region. Different subject areas had similar growth trends, and low growth is predicted to continue.

A similar method is used to predict the numbers of teacher leavers and not appropriately certified teachers. These predictions are calculated by multiplying the predicted number of teachers by intermediate statistics: teacher attrition rates and not appropriately certified teacher rates. This calculation is performed for each combination of region, grade level, and subject area.

Figure 11 provides an example of historical and intermediate statistics for not appropriately certified high school teachers in selected Missouri regions. The regions were selected to show the variability in rates by region, in both magnitude and change over time. The section below on validating intermediate statistics describes how the calculation method for these intermediate statistics is selected.

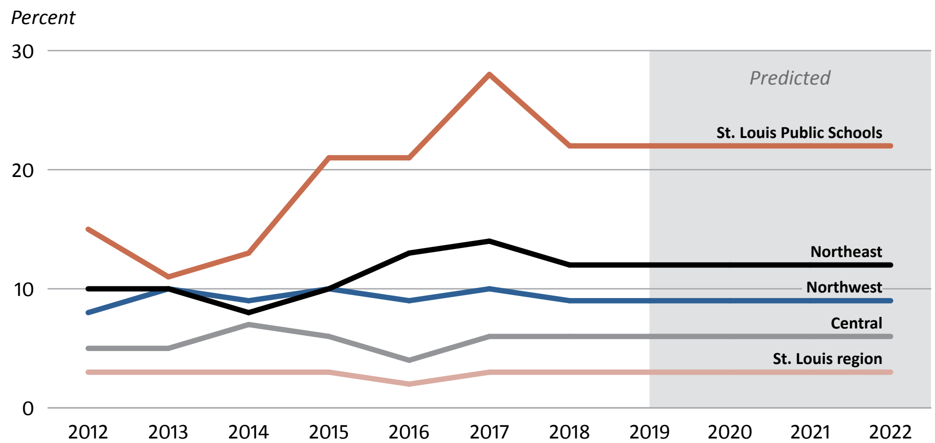
The validation process is intended to select the intermediate statistic calculation method. As in the example in figure 11, the intermediate statistics (teacher–student ratio, teacher attrition rate, and not appropriately certified teacher rate) are constant throughout the four years of the predictions. This does not mean that the future values of intermediate statistics are expected to be constant, only that constant prediction is the most likely result for a given year. It also does not mean that all predicted outputs are constant. Because predicted student enrollment counts can change over time, the model output predictions can also change, as in the example in figure 12.

**Figure 10. Historical and predicted numbers of employed middle school teachers in selected subject areas in the Kansas City region of Missouri, 2012–22**




Source: Authors' construction using data from the Missouri Department of Elementary and Secondary Education.

**Figure 11. Historical and predicted rates of not appropriately certified high school teachers in selected Missouri regions, 2012–22**




Source: Authors' construction using data from the Missouri Department of Elementary and Secondary Education.

In the Missouri model not appropriately certified rates were used as an indicator of teacher shortages.[2] The data show differences in historical and predicted rates by region. For example, teacher shortages are and will continue to be most severe in St. Louis City Schools. The variation across regions suggest that different policies or practices might be considered to address teacher shortages in different regions.

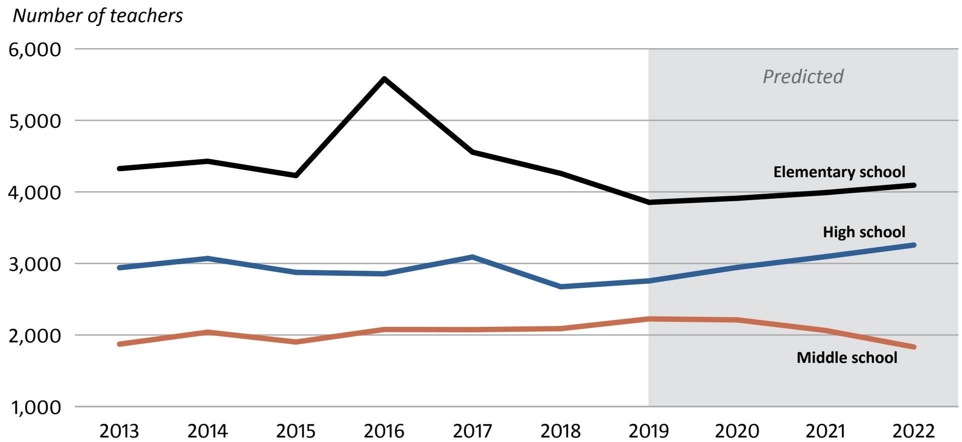
The model predicts new teacher hires by using other predictions from the model. Specifically, the number of new teacher hires ($NT_1$) is the predicted number of employed teachers for a given year ($T_1$) minus the predicted number of teachers retained from the prior year. The number of teachers retained from the prior year can be calculated by subtracting the number of teacher leavers ($L_0$) from the number of employed teachers in the prior year ($T_0$):

$$NT_1 = T_1 - (T_0 - L_0)$$

This method was used to predict the number of new teacher hires needed in Missouri from 2019 through 2022 for each combination of region, grade level, and subject area. Figure 12 provides an example, showing historical and predicted new teacher hires by grade level in Missouri. The number of new teacher hires is predicted to decrease among middle school teachers, reflecting the decline in elementary enrollment from 2016 to 2019.

2. As noted earlier, this indicator may underestimate teacher shortages. Because the model's shortage indictor is based on positions filled with not appropriately certified teachers, they do not include positions for which no teachers are available to teach, regardless of qualifications.

**Figure 12. Historical and predicted numbers of new teacher hires by grade level in Missouri, 2013–22**




Source: Authors' construction using data from the Missouri Department of Elementary and Secondary Education.

# Validating intermediate statistics

Because intermediate statistics are central to the teacher predictor model's predictions, it is important to select the most valid method to calculate the four intermediate statistics (student enrollment progress rate, teacher–student ratio, teacher attrition rate, and not appropriately certified teacher rate). The validation process involves using several methods to make predictions for each intermediate statistic based on early years in the observed data and then comparing the predictions with observed data in subsequent years. The method that most accurately describes actual trends is assumed to make the most accurate future predictions.

The Missouri Department of Elementary and Secondary Education considered four methods that are commonly used in teacher predictor models (Armstrong et al., 2011: Barro, 1992; Lindsay et al., 2009):

- The most recent year of observed data (most recent year).
- The average over the most recent two-year period (two-year average).
- The average over the most recent four-year period (four-year average).
- A straight-line trend that uses ordinary least squares regression and four years of observed data to make a one-year projection (straight-line trend).

In the Missouri model the predictions were made using data for 2012 through 2016 and compared with observed data for 2017 and 2018.

Each of the three steps of the validation process is presented below, along with an example from the Missouri model.

## Step 1: Making predictions using all four methods

Predictions were made using all four methods for each intermediate statistic. Table 2 provides an example, showing predicted teacher attrition rates in Missouri for 2017 and 2018 that were calculated using the four methods.

**Table 2. Predicted teacher attrition rates in Missouri, calculated using different methods**

| Predication method | Attrition rate |
|---|---|
| Most recent year | 0.150 |
| Four-year average | 0.149 |
| Two-year average | 0.147 |
| Straight-line trend | 0.146 |

Source: Authors' construction using data from the Missouri Department of Elementary and Secondary Education.

## Step 2: Comparing the predictions with actual data

The predictions from step 1 are then compared with actual data in order to identify the most valid method. Two metrics for comparison are used: the average percentage error and the mean absolute percentage error (Levin et al., 2015; Minnesota Department of Education, 2015). These metrics are calculated at the state and region levels for each intermediate statistic.

For each prediction (for example, each grade level for student enrollment predictions and each subject area for teacher predictions), the first step is to calculate the difference between predicted and actual rates. These differences are calculated within each grade level for each region. In the Missouri model the predictions in table 2 were subtracted from actual data for 2017 and 2018 to produce the differences in table 3. These differences are used to calculate the average percentage error. The mean absolute percentage error uses the absolute values of these differences.

**Table 3. Examples of differences between predicted and observed attrition rates for high school math teachers in the Central region of Missouri**

| | | 2017 | | | 2018 | | |
|---|---|---|---|---|---|---|---|
| Predication method | Predicted value | Observed value | Difference | Absolute value of the difference | Observed value | Difference | Absolute value of the difference |
| Most recent year | 0.150 | 0.145 | −0.005 | 0.005 | 0.148 | −0.002 | 0.002 |
| Four-year average | 0.149 | 0.145 | −0.004 | 0.004 | 0.148 | −0.001 | 0.001 |
| Two-year average | 0.147 | 0.145 | −0.002 | 0.002 | 0.148 | 0.001 | 0.001 |
| Straight-line trend | 0.146 | 0.145 | −0.001 | 0.001 | 0.148 | 0.002 | 0.002 |

Source: Authors' construction using data from the Missouri Department of Elementary and Secondary Education.

The average percentage error is used to determine whether predictions are biased in a certain direction. Average percentage error values can be positive or negative. To calculate the average percentage error, the positive and negative differences are summed across all the subject areas and regions and then divided by the sum of the observed values for the intermediate statistics to create a percentage difference. When summed, positive and negative differences can cancel each other out, but consistent differences in a positive or negative direction will become apparent.

The mean absolute percentage error is used to describe the size of the errors in the predictions. The process for calculating the mean absolute percentage error is similar to the process for calculating the average percentage error, except that the absolute values of the differences between predicted and actual values are summed across all the subject areas and regions and then divided by the sum of the observed values. The absolute values of the differences do not cancel each other out, creating a measure of the total size of the errors in the predictions. Thus, mean absolute percentage error values are always positive. For both the average percentage error and the mean absolute percentage error, the smaller the value, the more accurate the prediction.

In the Missouri model the average percentage error and the mean absolute percentage error were calculated separately for each grade level. They were also calculated for the state as a whole and for each region. Thus, both errors were calculated for each of the four methods (most recent year, four-year average, two-year average, and straight-line trend) at six grade-level (elementary school, middle school, and high school) and geographic (statewide and region) combinations (table 4).

**Table 4. Mean absolute percentage errors and average percentage errors used to validate methods for calculating teacher attrition rates in Missouri**

| | Elementary school | | Middle school | | High school | |
|---|---|---|---|---|---|---|
| Validity metric and prediction method | Statewide total difference | Regional difference | Statewide total difference | Regional difference | Statewide total difference | Regional difference |
| *Average percentage error* | | | | | | |
| Most recent year | −0.15 | −0.22 | 0.28 | 0.01 | −0.07 | −0.16 |
| Four-year average | −0.10 | −0.11 | −0.03 | 0.00 | −0.04 | −0.14 |
| Two-year average | −0.21 | −0.19 | −0.01 | −0.02 | −0.02 | −0.15 |
| Straight-line trend | −0.34 | −0.35 | −0.04 | −0.04 | −0.01 | −0.13 |
| *Mean absolute percentage error* | | | | | | |
| Most recent year | 0.50 | 0.67 | 0.45 | 0.50 | 0.16 | 0.51 |
| Four-year average | 0.12 | 0.40 | 0.08 | 0.28 | 0.13 | 0.34 |
| Two-year average | 0.25 | 0.48 | 0.07 | 0.27 | 0.14 | 0.39 |
| Straight-line trend | 0.43 | 0.66 | 0.12 | 0.43 | 0.15 | 0.68 |

Source: Authors' construction using data from the Missouri Department of Elementary and Secondary Education.

## Step 3: Identifying the most valid method for calculating each intermediate statistic

There are different processes for using the average percentage error and mean absolute percentage error to identify the most valid method for calculating each intermediate statistic. Missouri chose a scorecard approach to balance the importance of state and regional projections and to use multiple measures (Armstrong et al., 2011) The scorecard ranks the accuracy of each method on a scale from 1 (most accurate) to 4 (least accurate) for each grade level at the state and regional levels. In the Missouri model, for the elementary statewide total difference, the average percentage error and mean absolute percentage error for the four-year average method both rank as 1 (table 5).

**Table 5. Mean absolute percentage error ranks and average percentage error ranks used to validate methods for calculating teacher attrition rates in Missouri**

| Validity metric and prediction method | Elementary school | | | | Middle school | | | | High school | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Statewide total difference | | Regional difference | | Statewide total difference | | Regional difference | | Statewide total difference | | Regional difference | |
| | Value | Rank[a] | Value | Rank[a] | Value | Rank[a] | Value | Rank[a] | Value | Rank[a] | Value | Rank[a] |
| *Mean absolute percentage error* | | | | | | | | | | | | |
| Most recent year | –0.15 | 2 | –0.22 | 3 | 0.28 | 4 | 0.01 | 2 | –0.07 | 4 | –0.16 | 4 |
| Four-year average | –0.10 | 1 | –0.11 | 1 | –0.03 | 2 | 0.00 | 1 | –0.04 | 3 | –0.14 | 2 |
| Two-year average | –0.21 | 3 | –0.19 | 2 | –0.01 | 1 | –0.02 | 3 | –0.02 | 2 | –0.15 | 3 |
| Straight-line trend | –0.34 | 4 | –0.35 | 4 | –0.04 | 3 | –0.04 | 4 | –0.01 | 1 | –0.13 | 1 |
| *Average percentage error* | | | | | | | | | | | | |
| Most recent year | 0.50 | 4 | 0.67 | 4 | 0.45 | 4 | 0.50 | 4 | 0.16 | 4 | 0.51 | 3 |
| Four-year average | 0.12 | 1 | 0.40 | 1 | 0.08 | 2 | 0.28 | 2 | 0.13 | 1 | 0.34 | 1 |
| Two-year average | 0.25 | 2 | 0.48 | 2 | 0.07 | 1 | 0.27 | 1 | 0.14 | 2 | 0.39 | 2 |
| Straight-line trend | 0.43 | 3 | 0.66 | 3 | 0.12 | 3 | 0.43 | 3 | 0.15 | 3 | 0.68 | 4 |

a. On a scale from 1, most accurate, to 4, least accurate.

Source: Authors' construction using data from the Missouri Department of Elementary and Secondary Education.

Next, the 12 ranks for each method are summed, resulting in a final score for each method. The method with the lowest score is the most valid. Because the final score is simply a sum of the six ranks, this approach equally weights the different predictions (state, regional, and grade levels). For example, the four-year average method sums to a final score of 18. The final score for the four-year average method was the lowest of the four methods, so it was determined to be the most valid.[3]

3. Because the scoring process sums only ranks, information about the relative size of errors is lost. This approach may mask the magnitude of the differences between predicted and actual values (for example, a small mean absolute percentage error and a very large mean absolute percentage error might get the same score). Another option is to sum the average percentage errors and mean absolute percentage errors across the grade-level and geographic area combinations. This gives greater weight to large errors and could result in large errors in a small number of predictions driving the selection of the method. In addition, summing average percentage errors minimizes errors when two predictions have errors in different directions.

**Table 6. Scoring mean absolute percentage error ranks and average percentage error ranks for each method of calculating teacher attrition rates in Missouri**

| Validity metric and prediction method | Elementary school: Statewide total difference | Elementary school: Regional difference | Middle school: Statewide total difference | Middle school: Regional difference | High school: Statewide total difference | High school: Regional difference | Final score |
|---|---|---|---|---|---|---|---|
| *Mean absolute percentage error* | | | | | | | |
| Most recent year | 2 | 3 | 4 | 2 | 4 | 4 | |
| Four-year average | 1 | 1 | 2 | 1 | 3 | 2 | |
| Two-year average | 3 | 2 | 1 | 3 | 2 | 3 | |
| Straight-line trend | 4 | 4 | 3 | 4 | 1 | 1 | |
| *Average percentage error* | | | | | | | |
| Most recent year | 4 | 4 | 4 | 4 | 4 | 3 | |
| Four-year average | 1 | 1 | 2 | 2 | 1 | 1 | |
| Two-year average | 2 | 2 | 1 | 1 | 2 | 2 | |
| Straight-line trend | 3 | 3 | 3 | 3 | 3 | 4 | |
| *Scoring* | | | | | | | |
| Most recent year | 6 | 7 | 8 | 6 | 8 | 7 | 42 |
| Four-year average | 2 | 2 | 4 | 3 | 4 | 3 | 18 |
| Two-year average | 5 | 4 | 2 | 4 | 4 | 5 | 24 |
| Straight-line trend | 7 | 7 | 6 | 7 | 4 | 5 | 36 |

Source: Authors' construction using data from the Missouri Department of Elementary and Secondary Education.

Table 7 shows the method identified as the most valid for each intermediate statistic in Missouri's teacher predictor model. Users can apply a similar process to validate the calculation methods in their own model. They may find that their results differ from those for Missouri.

**Table 7. Calculation methods identified as the most valid for each intermediate statistic in Missouri's teacher predictor model**

| Intermediate statistic | Most valid method |
|---|---|
| *Student enrollment progression rates* | |
| Grade progression (K–grade 11) | Most recent year |
| Kindergarten enrollment | Two-year average |
| Prekindergarten enrollment | Straight-line trend |
| Teacher–student ratio | Four-year average |
| Teacher attrition rate | Most recent year |
| Not appropriately certified teacher rate | Four-year average |

Source: Authors' analysis using data from the Missouri Department of Elementary and Secondary Education

# Key considerations when developing a teacher predictor model

This section describes four topics that must be considered when developing a teacher predictor model. For each decision the benefits and drawbacks of different options are discussed, and the choices the Missouri Department of Elementary and Secondary Education made when developing its model are presented.

## Selecting the number of years to include in predictions

The number of years for which predictions are made must balance the needs of decisionmakers with the fact that the accuracy of predictions declines the farther they are in the future (Levin et al., 2015). Review of validation data (described in more detail below) indicates that mean absolute percentage errors increase by about 1 percentage point each year.

In the Missouri model the number of employed teachers was predicted for four years, a period that the Missouri Department of Elementary and Secondary Education felt was appropriate and that reflected the longest time frame for a decision based on these data. The department felt that four years is about the time it would take for an undergraduate student to select an education-related major in an identified area of need or for a person with a bachelor's degree to select a teacher preparation program.

## Accounting for teachers with multiple assignments

Some teachers have multiple teaching assignments, resulting in multiple records for those teachers in state administrative data systems. Each record typically describes one assignment and its associated subject area and grade level as well as the amount of time associated with the assignment (full-time equivalent). How to account for this data structure is a key decision point in designing and implementing a teacher predictor model. There are two primary options: include each individual teaching assignment in the analysis, using the full-time equivalent to account for the portion of a teacher's time that each assignment takes up, or count each individual teacher only once in the analysis, using information about his or her primary teaching assignment. Each option has benefits and drawbacks.

Although the full-time equivalent is a relevant unit in the K–12 school system, particularly for budgeting purposes, individual teachers are a more relevant unit of analysis for a teacher predictor model. The information from the model can aid potential teachers in making choices about topics to study, teacher preparation programs in developing programming for potential new teachers, and districts in recruiting and retaining teachers. Individual teachers are the common unit for these audiences. But full-time equivalents might be more precise when many teachers have multiple teaching assignments. The Missouri Department of Elementary and Secondary Education felt that the benefit of this additional precision did not outweigh the need for results that were useful to multiple audiences.

In some state administrative data systems, all required data may not be available in the full-time equivalent metric. For example, in a Massachusetts study, data on new and retained teachers were available only as counts of teachers, not full-time equivalents (Levin et al., 2015). As a result, the researchers had to use the counts to estimate the full-time equivalents for the new teachers and retained teachers. Several previous studies have based analyses on counts of teachers instead of full-time equivalents, likely because counts provide more interpretable findings (for example, Berg-Jacobson & Levin, 2015).

For an analysis based on counts of individual teachers, a teacher predictor model requires that there be only one record for each teacher, with a unique identifier and information about the teacher's primary district, grade level, and subject area. To guide the process of converting teacher records from multiple records per teacher to one record per teacher, users should develop rules for determining a teacher's primary assignment. These rules should reflect available data and common practices for imputing information that describes teaching assignments. For example, if a state administrative data system includes only the number of subject-area and grade-level assignments (not full-time equivalents) for each teacher, the rules could determine a teacher's primary assignment based on the number of assignments. Or if a state administrative data system includes full-time equivalent data, but the full-time equivalents for traditional elementary school teachers are distributed equally among math, English language arts, science, and fine arts, the rules could determine that the primary assignment for all teachers in elementary schools with equal full-time equivalents across these subject areas is elementary education.

In the Missouri model primary teaching assignments were determined based on the districts, grade levels, and subject areas in which teachers were assigned the majority of their time. The Missouri Department of Elementary and Secondary Education's goal was to use the available information to accurately describe teachers' assignments without undercounting shortage areas. A standard procedure was developed to address each case in which teachers spent equal amounts of time in grade levels or subject areas. For example, the primary assignment for a teacher with equal proportions of time allocated across different subject areas was the subject area for which the teacher was not appropriately certified. If the teacher was appropriately certified to teach both subject areas, the teacher's primary assignment was the subject with more historical evidence of a shortage. This procedure was intended to provide a complete picture of teachers working in subject areas facing shortages in terms of both appropriately and not appropriately certified teachers. This procedure might underestimate or overestimate the number of appropriately certified teachers in particular grade levels or subject areas.

## Determining the number of years of historical data to include

A teacher predictor model allows users to determine the number of years of historical data to include. Previous research using similar models varies in the number of years of data included (Berg-Jacobsen & Levin, 2015; Levin et al., 2015). Building a model requires at least three years of data, and validating a model requires at least five years of data. The validation process involves making predictions based on earlier years within the available data and then comparing those predictions with actual data in subsequent years.

When determining the number of years to include, users should consider two key issues. First, users should ensure that data are reliable and valid—that is, that data are recorded in a consistent manner for all years that are included in the model. Second, users should ensure that data from time periods with substantial differences in economic or policy conditions are carefully evaluated, as the model assumes that historical conditions will continue into the future. During periods of substantial economic change, model parameters should be reviewed and updated as needed. In the Missouri model, the Department of Elementary and Secondary Education chose to use data from the eight most recent academic years, reflecting all available data since changes affecting data reliability had been made to the state data system.

## Determining which subgroups of teachers to examine

A teacher predictor model may be used to make predictions for a variety of teacher subgroups, defined by variables such as subject area, grade level, geographic region, race/ethnicity, and age. For example, a model may predict a future shortage of high school special education teachers in the southeast region of a state.

In the Missouri model predictions describe the number of employed teachers, teacher leavers, new hires, and not appropriately certified teachers in all of the following:

- 21 subject areas, as well as totals of all subject areas.
- 3 grade levels (elementary school, middle school, and high school).
- 11 regions of Missouri as well as the state as a whole.

As in any statistical model, the number of subgroup variables that can be considered simultaneously depends to some degree on the total number of cases in the data being analyzed. Increasing the number of variables that are examined simultaneously reduces the precision of the model. For that reason, users should consider including variables that define the teacher subgroups that are most important to their local contexts and that are based on how they anticipate the model results will be used.

The variables that are most important to include will likely vary across local contexts. For example, in a study in Massachusetts, projections related to teacher race/ethnicity, teacher age, and the rurality of the schools in which teachers taught were the primary interests (Levin et al., 2015). In the Missouri model the Department of Elementary and Secondary Education examined multiple subject-area assignments in science, resulting in relatively small cell sizes. The department identified physics, chemistry, earth science, and biology because those subjects reflected different coursework during preparation. The department focused on grade levels, subject areas, and regions in order to inform potential teachers of future opportunities and to advise teacher preparation programs of possible areas of focus.

# Conclusions

This resource describes how to develop and validate a teacher predictor model, using Missouri's model as an example. The Missouri Department of Elementary and Secondary Education developed the teacher predictor model to identify and predict teacher shortages. Department staff can implement the model with available software, using student enrollment and teacher workforce data that are available in state administrative data systems. The model described here improves on shortage estimates based only on the numbers of teachers who are not appropriately certified, providing multiple pieces of information about the teacher workforce: size, attrition, number of new hires, and shortages. This information is provided for each region in the state and for multiple subject areas. To encourage buy-in for using the model results, the model is designed to be easily understood by administrators and other leaders.

This resource identifies key decisions that must be made when developing a teacher predictor model. These decisions should be guided by the goals and purposes of the model, which are likely to vary across states. Identifying the goals and priorities can support decisionmaking throughout the model development process. To develop a viable model for Missouri, the Department of Elementary and Secondary Education focused on ensuring data quality, including reliability and validity, and on providing useful information to stakeholders concerned about the teacher preparation pipeline.

Information from the Missouri model is being used to support a collaborative effort among Missouri Department of Elementary and Secondary Education staff, school district staff, and representatives of teacher preparation programs. This collaboration continues as stakeholders consider how to incorporate more information into the model, including information to address equity concerns raised by shortages, district survey data on shortages and vacancies, and contextual data on sources of supply, such as hiring patterns from local teacher preparation programs. This collaboration aims to help stakeholders proactively identify strategies to address current teacher shortages and anticipated future shortage areas and evaluate the future progress of those strategies.

# Appendix A. Teacher predictor model calculations

This appendix presents sample calculations for each step of the teacher predictor model. As described in the main text, the model uses student enrollment data as an input to calculate annual student enrollment progression rates. These calculations are made for every year of observed data. Table A1 presents sample calculations, using student enrollment data, of annual progression rates for grades 1–12.

Next, using the results of the student enrollment progression rates, four methods are considered for calculating the progression rate that will be used to make student enrollment predictions, called the validated progression rate. This process is described in the Validating intermediate statistics section of the main text. The predictions made with each method are compared with actual data to determine which approach is the most valid. The validated progression rate is then used to predict student enrollment. Table A2 presents sample calculations for this step. As described in the Predicting student enrollment section of the main text, progression rates are not calculated for prekindergarten and kindergarten. Instead, prior enrollment is used. Using the process described in the Validating intermediate statistics section of the main text, a validated enrollment intermediate statistic is identified for each of these grade levels.

**Table A1. Sample calculations of student enrollment progression rates from 2017 to 2018 for grades 1–12**

| Grade | Calculation for student enrollment progression rate from 2017 to 2018 |
|---|---|
| 12 | 2018 enrollment in grade 12 ÷ 2017 enrollment in grade 11 |
| 11 | 2018 enrollment in grade 11 ÷ 2017 enrollment in grade 10 |
| 10 | 2018 enrollment in grade 10 ÷ 2017 enrollment in grade 9 |
| 9 | 2018 enrollment in grade 9 ÷ 2017 enrollment in grade 8 |
| 8 | 2018 enrollment in grade 8 ÷ 2017 enrollment in grade 7 |
| 7 | 2018 enrollment in grade 7 ÷ 2017 enrollment in grade 6 |
| 6 | 2018 enrollment in grade 6 ÷ 2017 enrollment in grade 5 |
| 5 | 2018 enrollment in grade 5 ÷ 2017 enrollment in grade 4 |
| 4 | 2018 enrollment in grade 4 ÷ 2017 enrollment in grade 3 |
| 3 | 2018 enrollment in grade 3 ÷ 2017 enrollment in grade 2 |
| 2 | 2018 enrollment in grade 2 ÷ 2017 enrollment in grade 1 |
| 1 | 2018 enrollment in grade 1 ÷ 2017 enrollment in kindergarten |

Source: Authors' construction.

**Table A2. Sample calculations using validated progression rates from 2017 to 2018 to predict student enrollment in 2019**

| Grade | Calculation for predicted student enrollment in 2019 |
|---|---|
| 12 | 2018 enrollment in grade 11 × grade 12 validated progression rate |
| 11 | 2018 enrollment in grade 10 × grade 11 validated progression rate |
| 10 | 2018 enrollment in grade 9 × grade 10 validated progression rate |
| 9 | 2018 enrollment in grade 8 × grade 9 validated progression rate |
| 8 | 2018 enrollment in grade 7 × grade 8 validated progression rate |
| 7 | 2018 enrollment in grade 6 × grade 7 validated progression rate |
| 6 | 2018 enrollment in grade 5 × grade 6 validated progression rate |
| 5 | 2018 enrollment in grade 4 × grade 5 validated progression rate |
| 4 | 2018 enrollment in grade 3 × grade 4 validated progression rate |
| 3 | 2018 enrollment in grade 2 × grade 3 validated progression rate |
| 2 | 2018 enrollment in grade 1 × grade 2 validated progression rate |
| 1 | 2018 enrollment in kindergarten × grade 1 validated progression rate |
| Kindergarten | Validated kindergarten enrollment |
| Prekindergarten | Validated prekindergarten enrollment |

Note: Initial calculations of progression rates are in table A1. After the initial calculations are made, the process described in the Validating intermediate statistics section of the main text is followed to arrive at the validated progression rate that is used in the calculations in this table.

Source: Authors' construction.

The teacher–student ratio is used to predict the number of employed teachers. Teacher–student ratios are calculated for every region, grade level (elementary school, middle school, and high school), and subject area. Table A3 presents sample calculations for high school teacher–student ratios for selected subject areas in the Northeast region of Missouri. After teacher–student ratios are calculated, the process described in the Validating intermediate statistics section of the main text is followed to arrive at the validated teacher–student ratios that are used to make predictions.

To predict the number of employed teachers in future years, the validated teacher–student ratios for each region, grade level, and subject area are multiplied by predicted student enrollment. Table A4 presents sample calculations for predicted employed high school teacher counts in 2019 for selected subject areas in the Northeast region of Missouri.

**Table A3. Sample calculations of high school teacher–student ratios for selected subject areas in the Northeast region of Missouri**

| Subject area | Calculation for teacher–student ratio |
|---|---|
| Agriculture | Count of agriculture teachers in the Northeast region ÷ high school enrollment count in the Northeast region |
| Art | Count of art teachers in the Northeast region ÷ high school enrollment count in the Northeast region |
| Biology | Count of biology teachers in the Northeast region ÷ high school enrollment count in the Northeast region |
| Business | Count of business teachers in the Northeast region ÷ high school enrollment count in the Northeast region |
| Chemistry | Count of chemistry teachers in the Northeast region ÷ high school enrollment count in the Northeast region |

Source: Authors' construction.

**Table A4. Sample calculations using validated teacher–student ratios to predict employed high school teacher counts in 2019 for selected subject areas in the Northeast region of Missouri**

| Subject area | Calculation for predicted teacher counts |
|---|---|
| Agriculture | 2019 predicted high school enrollment in the Northeast region × validated teacher–student ratio for high school agriculture teachers in the Northeast region |
| Art | 2019 predicted high school enrollment in the Northeast region × validated teacher–student ratio for high school art teachers in the Northeast region |
| Biology | 2019 predicted high school enrollment in the Northeast region × validated teacher–student ratio for high school biology teachers in the Northeast region |
| Business | 2019 predicted high school enrollment in the Northeast region × validated teacher–student ratio for high school business teachers in the Northeast region |
| Chemistry | 2019 predicted high school enrollment in the Northeast region × validated teacher–student ratio for high school chemistry teachers in the Northeast region |

Note: Calculations of predicted high school enrollment are in table A2. Calculations are initially made for each grade and then summed across grades 9–12 to arrive at total high school predicted enrollment. Initial calculations of teacher–student ratios are in table A3. After the initial calculations are made, the process described in the Validating intermediate statistics section of the main text is followed to arrive at the validated teacher–student ratio that is used in the calculations in this table.

Source: Authors' construction.

Teacher attrition rates are calculated for each academic year by dividing the number of teachers who left districts by the total number of teachers. Because calculating attrition rates requires two subsequent years of data, the rates are not calculated for the final year of observed data. Teacher attrition rates are calculated for each region, grade level (elementary school, middle school, and high school), and subject area. Table A5 presents sample calculations for high school teacher attrition rates from 2017 to 2018 for selected subject areas in the Northeast region of Missouri. Prior to performing these calculations, teacher employment data for 2017 and 2018 were compared to determine how many teachers were leavers. Leavers are defined as teachers who no longer taught in the same district in 2018. After teacher attrition rates are calculated, the process described in the Validating intermediate statistics section of the main text is followed to arrive at the validated teacher attrition rates that are used to make predictions.

To predict the number of teacher leavers in future years, the predicted number of teachers is multiplied by the validated teacher attrition rate. These calculations are made for every region, grade level (elementary school, middle school, and high school), and subject area. Table A6 presents sample calculations for predicted counts of high school teacher leavers in 2019 for selected subject areas in the Northeast region.

**Table A5. Sample calculations of high school teacher attrition rates from 2017 to 2018 for selected subject areas in the Northeast region of Missouri**

| Subject area | Calculation for teacher attrition rates |
|---|---|
| Agriculture | 2018 count of agriculture teacher leavers in the Northeast region ÷ 2017 count of agriculture teachers in the Northeast region |
| Art | 2018 count of art teacher leavers in the Northeast region ÷ 2017 count of art teachers in the Northeast region |
| Biology | 2018 count of biology teacher leavers in the Northeast region ÷ 2017 count of biology teachers in the Northeast region |
| Business | 2018 count of business teacher leavers in the Northeast region ÷ 2017 count of business teachers in the Northeast region |
| Chemistry | 2018 count of chemistry teacher leavers in the Northeast region ÷ 2017 count of chemistry teachers in the Northeast region |

Note: Teacher leavers in 2018 are teachers who were employed in 2017 but were no longer employed in the same district in 2018.

Source: Authors' construction.

**Table A6. Sample calculations using validated teacher attrition rates to predict counts of high school teacher leavers in 2019 for selected subject areas in the Northeast region of Missouri**

| Subject area | Calculation for predicted count of teacher leavers |
|---|---|
| Agriculture | 2019 predicted count of high school agriculture teachers in the Northeast region × validated teacher attrition rate for high school agriculture teachers in the Northeast region |
| Art | 2019 predicted count of high school art teachers in the Northeast region × validated teacher attrition rate for high school art teachers in the Northeast region |
| Biology | 2019 predicted count of high school biology teachers in the Northeast region × validated teacher attrition rate for high school biology teachers in the Northeast region |
| Business | 2019 predicted count of high school business teachers in the Northeast region × validated teacher attrition rate for high school business teachers in the Northeast region |
| Chemistry | 2019 predicted count of high school chemistry teachers in the Northeast region × validated teacher attrition rate for high school chemistry teachers in the Northeast region |

Note: Calculations of predicted teacher counts are in table A4. Initial calculations of teacher attrition rates are in table A5. After the initial calculations are made, the process described in the Validating intermediate statistics section of the main text is followed to arrive at the validated teacher attrition rate that is used in the calculations illustrated in this table.

Source: Authors' construction.

Not appropriately certified teacher rates are calculated for each academic year by dividing the number of teachers in a given year who are not appropriately certified by the total number of teachers. These calculations are made for each region and grade level (elementary school, middle school, and high school). Table A7 presents sample calculations for not appropriately certified high school teacher rates in 2018 for selected subject areas in the Northeast region of Missouri. After not appropriately certified teacher rates are calculated, the process described in the Validating intermediate statistics section of the main text is followed to arrive at the validated not appropriately certified teacher rates that are used to make predictions.

To predict the number of not appropriately certified teachers in future years, the predicted number of teachers is multiplied by the validated not appropriately certified teacher rate. These calculations are made for every region, grade level (elementary school, middle school, and high school), and subject area. Table A8 presents sample calculations for predicted counts of not appropriately certified high school teachers in 2019 for selected subject areas in the Northeast region of Missouri.

**Table A7. Sample calculations of not appropriately certified high school teacher rates in 2018 for selected subject areas in the Northeast region of Missouri**

| Subject area | Calculation for not appropriately certified teacher rate |
|---|---|
| Agriculture | 2018 count of not appropriately certified agriculture teachers in the Northeast region ÷ 2018 count of agriculture teachers in the Northeast region |
| Art | 2018 count of not appropriately certified art teachers in the Northeast region ÷ 2018 count of art teachers in the Northeast region |
| Biology | 2018 count of not appropriately certified biology teachers in the Northeast region ÷ 2018 count of biology teachers in the Northeast region |
| Business | 2018 count of not appropriately certified business teachers in the Northeast region ÷ 2018 count of business teachers in the Northeast region |
| Chemistry | 2018 count of not appropriately certified chemistry teachers in the Northeast region ÷ 2018 count of chemistry teachers in the Northeast region |

Source: Authors' construction.

**Table A8. Sample calculations using validated not appropriately certified teacher rates to predict counts of not appropriately certified high school teachers in 2019 for selected subject areas in the Northeast region of Missouri**

| Subject area | Calculation for predicted count of not appropriately certified teachers |
|---|---|
| Agriculture | 2019 predicted count of high school agriculture teachers in the Northeast region × validated not appropriately certified rate for high school agriculture teachers in the Northeast region |
| Art | 2019 predicted count of high school art teachers in the Northeast region × validated not appropriately certified rate for high school art teachers in the Northeast region |
| Biology | 2019 predicted count of high school biology teachers in the Northeast region × validated not appropriately certified rate for high school biology teachers in the Northeast region |
| Business | 2019 predicted count of high school business teachers in the Northeast region × validated not appropriately certified rate for high school business teachers in the Northeast region |
| Chemistry | 2019 predicted count of high school chemistry teachers in the Northeast region × validated not appropriately certified rate for high school chemistry teachers in the Northeast region |

Note: Calculations of predicted teacher counts are in table A4. Initial calculations of not appropriately certified rate are in table A7. After the initial calculations are made, the process described in the Validating intermediate statistics section of the resource is followed to arrive at the validated not appropriately certified rate that is used in the calculations in this table.

Source: Authors' construction.

The teacher predictor model predicts new teacher hires by using other predictions from the model. Specifically, the number of new teacher hires is the predicted number of employed teachers for a given year minus the predicted number of teachers retained from the prior year. These calculations are made for every region, grade level (elementary school, middle school, and high school), and subject area. Table A9 presents sample calculations for predicted counts of retained high school teachers in 2020 for selected subject areas in the Northeast region of Missouri.

Next, the predicted count of retained teachers is subtracted from the predicted counts of teachers to calculate the predicted number of new hires. These calculations are made for every region, grade level (elementary school, middle school, and high school), and subject area. Table A10 presents sample calculations for predicted counts of high school teacher new hires in 2020 for selected subject areas in the Northeast region of Missouri.

**Table A9. Sample calculations of predicted counts of retained high school teachers in 2020 for selected subject areas in the Northeast region of Missouri**

| Subject area | Calculation for predicted count of retained teachers |
|---|---|
| Agriculture | 2019 predicted count of high school agriculture teachers in the Northeast region – 2019 predicted count of agriculture teacher leavers in the Northeast region |
| Art | 2019 predicted count of high school art teachers in the Northeast region – 2019 predicted count of art teacher leavers in the Northeast region |
| Biology | 2019 predicted count of high school biology teachers in the Northeast region – 2019 predicted count of biology teacher leavers in the Northeast region |
| Business | 2019 predicted count of high school business teachers in the Northeast region – 2019 predicted count of business teacher leavers in the Northeast region |
| Chemistry | 2019 predicted count of high school chemistry teachers in the Northeast region – 2019 predicted count of chemistry teacher leavers in the Northeast region |

Note: Calculations for predicted teacher counts are in table A4. Calculations for predicted counts of teacher leavers are in table A6.

Source: Authors' construction.

**Table A10. Sample calculations of predicted counts of high school teacher new hires in 2020 for the Northeast region of Missouri**

| Subject area | Calculation for predicted count of new hires |
|---|---|
| Agriculture | 2020 predicted count of high school agriculture teachers in the Northeast region – 2020 predicted count of retained agriculture teachers in the Northeast region |
| Art | 2020 predicted count of high school art teachers in the Northeast region – 2020 predicted count of retained art teachers in the Northeast region |
| Biology | 2020 predicted count of high school biology teachers in the Northeast region – 2020 predicted count of retained biology teachers in the Northeast region |
| Business | 2020 predicted count of high school business teachers in the Northeast region – 2020 predicted count of retained business teachers in the Northeast region |
| Chemistry | 2020 predicted count of high school chemistry teachers in the Northeast region – 2020 predicted count of retained chemistry teachers in the Northeast region |

Note: Calculations of predicted teacher counts are in table A4. Calculations of predicted counts of retained teachers are in table A9.

Source: Authors' construction.

# References


Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, *25*(1), 95–135. https://doi.org/10.1086/508733.

Aldeman, C. (2018). *Educator supply and demand reports: Lessons for states* [PowerPoint]. Bellwether Education Partners. Retrieved March 28, 2020, from https://bellwethereducation.org/publication/teacher-supply-and-demand-how-states-track-shortage-areas.

Armstrong, J. S., Green, K. C., & Graefe, A. (2011). Forecasting principles. In M. Lovric (Ed.), *International encyclopedia of statistical science* (pp. 417–440). Springer-Verlag. Retrieved May 14, 2020, from http://repository.upenn.edu/marketing_papers/178

Barro, S. M. (1992). Models for projecting teacher supply, demand, and quality: An assessment of the state of the art. In E. E. Boe & D. M. Gilford (Eds.), *Teacher supply, demand, and quality: Policy issues, models, and data bases* (pp. 129–209). National Academies Press. Retrieved May 14, 2020, from https://www.nap.edu/catalog/2040/teacher-supply-demand-and-quality-policy-issues-models-and-data.

Berg-Jacobson, A., & Levin, J. (2015). *Oklahoma study of educator supply and demand: Trends and projections.* American Institutes for Research. https://eric.ed.gov/?id=ED578797.

Boe, E. E., & Gilford, D. M. (1992). Summary of conference proceedings. In E. E. Boe & D. M. Gilford (Eds.), *Teacher supply, demand, and quality: Policy issues, models, and data bases* (pp. 21–62). National Academies Press. Retrieved May 14, 2020, from https://www.nap.edu/catalog/2040/teacher-supply-demand-and-quality-policy-issues-models-and-data.

Borman, G. D., & Dowling, N. M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research, 78*(3), 367–409. https://eric.ed.gov/?id=EJ896364.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management, 24*(1), 113–132. https://eric.ed.gov/?id=EJ759320.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, *104*(9), 2593–2632. Retrieved March 28, 2020, from https://www.aeaweb.org/articles?id=10.1257/aer.104.9.2593

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *Journal of Human Resources, 45*(3), 655–681. https://eric.ed.gov/?id=EJ889247.

Education Commission of the States. (2019, October). *Has the state published state-specific teacher shortage data within the past five years?* Retrieved May 14, 2020, from http://ecs.force.com/mbdata/MBQuest2rtanw?Rep=TRR1903.

Folsom, J. S., Osborne-Lampkin, L., & Herrington, C. D. (2014). *Using administrative data for research: A companion guide to a descriptive analysis of the principal workforce in Florida schools* (REL 2015–049). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. https://eric.ed.gov/?id=ED551066.

Glazerman, S., & Max, J. (2011). *Do low-income students have equal access to the highest-performing teachers?* (NCEE No. 2011–4016). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://eric.ed.gov/?id=ED517967.

Ingersoll, R. M. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal, 38*(3), 499–534. Retrieved May 14, 2020, from https://repository.upenn.edu/gse_pubs/94/.

Isenberg, E., Max, J., Gleason, P., Johnson, M., Deutsch, J., & Hansen, M. (2016). *Are high-and low-income students taught by equally effective teachers?* (NCEE No. 2017–4009). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/pubs/20174008/.

Isenberg, E., Max, J., Gleason, P., Potamites, L., Santillano, R., Hock, H., & Hansen, M. (2013). *Access to effective teaching for disadvantaged students* (NCEE No. 2014–4001). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://eric.ed.gov/?id=ED544345.

Konstantopoulos, S., & Chung, V. (2011). The persistence of teacher effects in elementary grades. *American Educational Research Journal*, *48*(2), 361–386. https://eric.ed.gov/?id=EJ921703.

Levin, J., Berg-Jacobson, A., Atchison, D., Lee, K., & Vontsolos, E. (2015). *Massachusetts study of teacher supply and demand: Trends and projections*. American Institutes for Research. https://eric.ed.gov/?id=ED563815.

Lindsay, J., Wan, Y., Berg-Jacobson, A., Walston, J., & Redford, J. (2016). *Strategies for estimating teacher supply and demand using student and teacher data* (REL 2017–197). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. https://eric.ed.gov/?id=ED570977

Lindsay, J. J., Wan, Y., & Gossin-Wilson, W. (2009). *Methodologies used by Midwest Region states for studying teacher supply and demand* (REL 2008–No. 080). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. https://eric.ed.gov/?id=ED506631.

Malkus, N., Mulvaney Hoyer, K., & Sparks, D. (2015). *Teaching vacancies and difficult-to-staff teaching positions in public schools* (NCES No. 2015–065). U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. http://eric.ed.gov/?id=ED561224.

Meyer, S. J., Espel, E. V., Weston-Sementelli, J. L., & Serdiouk, M. (2019). *Teacher retention, mobility, and attrition in Colorado, Missouri, Nebraska, and South Dakota.* (REL 2019–001). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Central. https://eric.ed.gov/?id=ED593492.

Minnesota Department of Education. (2015). *Teacher supply and demand: Fiscal year 2015 report to the legislature*. https://eric.ed.gov/?id=ED585479

Missouri Department of Elementary and Secondary Education. (2015). *Ensure equitable access to excellent educators*. https://dese.mo.gov/sites/default/files/Educator-Equity-Plan-Missouri.pdf

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, *26*(3), 237–257. https://eric.ed.gov/?id=EJ727526.

Pennington McVey, K., & Trinidad, J. (2019). *Nuance in the noise: The complex reality of teacher shortages*. Bellwether Education Partners. https://eric.ed.gov/?id=ED596444.

Pettibone, T. J., & Bushan, L. (1990, November). *School district enrollment projections: A comparison of three methods*. Paper presented at the annual meeting of the Mid-South Educational Research Association. https://eric.ed.gov/?id=ED326560.

Reichardt, R. (2003). *Teacher supply and demand in the state of Colorado.* Mid-continent Research for Education and Learning. http://hermes.cde.state.co.us/drupal/islandora/object/co:2133/datastream/OBJ/download/Teacher_supply_and_demand_in_the_state_of_Colorado.pdf

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal, 50*(1), 4–36. https://eric.ed.gov/?id=EJ995828.

Sass, T. R., Hannaway, J., Xu, Z., Figlio, D. N., & Feng, L. (2012). Value added of teachers in high-poverty schools and lower poverty schools. *Journal of Urban Economics, 72*(2–3), 104–122. https://doi.org/10.1016/j.jue.2012.04.004.

Stewart, J., Rhoads, C., Serdiouk, M., Van Dine, D., Cherasaro, T., & Klute, M. (2019). *Associations between the qualifications of middle school Algebra I teachers and student math achievement* (REL 2020–005). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Central. https://eric.ed.gov/?id=ED598960.

Sutcher, L., Darling-Hammond, L., & Carver-Thomas, D. (2016). *A coming crisis in teaching? Teacher supply, demand, and shortages in the U.S.* Learning Policy Institute. Retrieved March 28, 2020, from https://learningpolicyinstitute.org/product/coming-crisis-teaching

Sweeney, S. H., & Middleton, E. J. (2005). Multiregional cohort enrolment projections: Matching methods to enrolment policies. *Population, Space and Place, 11*(5), 361–380. https://doi.org/10.1002/psp.379.

U.S. Department of Education, Office of Postsecondary Education. (2015). *Enrollment in teacher preparation programs*. http://eric.ed.gov/?id=ED576131

White, M. E., & Fong, A. B. (2008). *Trends in California teacher demand: a county and regional perspective* (REL 2008–No. 057). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West. https://eric.ed.gov/?id=ED502239.

**Acknowledgments**

This project would not have been possible without the leadership of Paul Katnik, associate commissioner of education at the Missouri Department of Elementary and Secondary Education, and the collaboration of staff at the department, with a special thank you to Buddy Alberson and Tim Wittman. The authors also acknowledge the contributions of Charles Harding and Douglas Gagnon at Marzano Research and Michael Allen and Michael Podgursky of the REL Central Technical Working Group.

REL 2021–052

December 2020

This resource 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.



Reichardt, R., Klute, M., Stewart, J., & Meyer, S. (2020). *An approach to using student and teacher data to understand and predict teacher shortages* (REL 2021–052). 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.