

**Study of the Distribution of
Teacher Effectiveness**

Part A

June 9, 2011



MATHEMATICA
Policy Research, Inc.

Contract Number:
ED-IES-10-C-0065

Mathematica Reference Number:
06860.170

Submitted to:
Institute of Education Sciences
U.S. Department of Education
555 New Jersey Avenue, NW
Washington, DC 20208
Project Officer: Lauren Angelo

Submitted by:
Mathematica Policy Research
600 Maryland Avenue, SW
Suite 550
Washington, DC 20024-2512
Telephone: (202) 484-9220
Facsimile: (202) 863-1763
Project Director: Philip Gleason

Study of the Distribution of Teacher Effectiveness

Part A

June 9, 2011



MATHEMATICA
Policy Research, Inc.

CONTENTS

SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT

PART A. JUSTIFICATION..... 1

- 1. Circumstances Necessitating the Collection of Information..... 1
 - a. Statement of Need to Study the Distribution of Teacher Effectiveness..... 1
 - b. Research Questions..... 2
 - c. Study Overview..... 2
 - d. Recruitment of Districts..... 3
 - e. Data Collection Plan and Study Timeline..... 5
- 2. Purposes and Uses of Data..... 6
- 3. Use of Technology to Reduce Burden..... 6
- 4. Efforts to Avoid Duplication..... 7
- 5. Methods to Minimize Burden on Small Entities..... 7
- 6. Consequences of Not Collecting Data..... 7
- 7. Special Circumstances..... 8
- 8. Federal Register Announcement and Consultation..... 8
 - a. Federal Register Announcement..... 8
 - b. Consultations Outside the Agency..... 8
 - c. Unresolved Issues..... 8
- 9. Payments or Gifts..... 8
- 10. Assurances of Confidentiality..... 8
- 11. Justification for Sensitive Questions..... 9
- 12. Estimates of Burden Hours..... 10
- 13. Estimates of Cost Burden to Respondents..... 11
- 14. Annualized Cost to the Federal Government..... 11
- 15. Reasons for Program Changes or Adjustments..... 11
- 16. Plans for Tabulation and Publication of Results..... 11
 - a. Tabulation Plans..... 11
 - i. Describing District Policies to Address the Distribution of Teacher Effectiveness..... 11

- ii. Measuring Teacher Effectiveness.....12
- iii. Measuring the Distribution of Teacher Effectiveness.....14
- iv. Analyzing the Distribution of Teacher Effectiveness Within
and Across Districts.....17
- b. Publication Plans.....18
- 17. Approval Not to Display the Expiration Date for OMB Approval.....19
- 18. Exception to the Certification Statement.....19
- REFERENCES.....20
- APPENDIX A: DISTRICT NOTIFICATION LETTER
- APPENDIX B: DISTRICT RECRUITMENT CONVERSATION GUIDE
- APPENDIX C: DISTRICT DATA REQUEST MEMO
- APPENDIX D: DISTRICT STAFF INTERVIEW PROTOCOL
- APPENDIX E: STUDY SUMMARY
- APPENDIX F: CONFIDENTIALITY PLEDGE

TABLES

1	Data Collection Plan.....	5
2	Estimated Response Time for Data Collection.....	10

SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT

This submission is a request for approval of recruitment and data collection activities that will be used to support the Study of the Distribution of Teacher Effectiveness (DTE). The study is being funded by the Institute of Education Sciences (IES), U.S. Department of Education (ED), and is being implemented by Mathematica Policy Research and its subcontractor, the Urban Institute.

The goal of the study is to examine the distribution of teacher effectiveness in up to 30 school districts and document changes in the distribution over time. We will use a value-added analysis to measure teacher effectiveness and compare the average value-added scores of teachers of disadvantaged and non-disadvantaged students, with eligibility for free and reduced-price lunch (FRL) as the primary measure of student disadvantage. The study will provide information on the distribution of teacher effectiveness in participating districts for two baseline years and three follow-up years. Interviews with district staff will provide information on district strategies to promote an equitable distribution of teacher effectiveness, allowing us to analyze the relationship between district policies and the distribution. The study will also explore the relationship between teacher mobility and the distribution of teacher effectiveness.

This submission requests approval to recruit districts for the study, collect student records and teacher personnel data, and conduct telephone interviews with staff in participating districts.

PART A. JUSTIFICATION

1. Circumstances Necessitating the Collection of Information

a. Statement of Need to Study the Distribution of Teacher Effectiveness

The specific legislation authorizing this data collection is Title II, Part A of the Elementary and Secondary Education Act (ESEA), Section 2121-2123, as amended by the No Child Left Behind Act (NCLB) (20 USC 6621-6623). Title II, Part A of ESEA provides funding to states to prepare, train, and recruit high-quality teachers to help school districts ensure that all students have effective teachers. One allowable use of Title II, Part A funds is to develop and implement initiatives to recruit and retain high-quality teachers in hard-to-staff schools. Part F, Section 9601 of ESEA permits program funds to be used to evaluate activities authorized under the act.

The focus of federal policy has shifted from targeting the distribution of highly qualified teachers to the distribution of highly effective teachers. While ESEA requires that states and districts ensure an equitable distribution of highly qualified teachers (HQT), the American Recovery and Reinvestment Act (ARRA) and other federal initiatives are aimed, at least in part, at

improving the distribution of effective teachers. The ARRA-funded Race to the Top (RTT) grants require that states implement strategies to ensure that students in high poverty schools and schools with large minority populations have equal access to effective teachers. The Teacher Incentive Fund (TIF) grants, also funded by ARRA, promote an equitable distribution by supporting incentives that reward teachers based on their performance and attract effective teachers into high-need schools.

Although there is growing concern that the nation's most talented teachers are not working in the most disadvantaged schools, there is limited evidence on the distribution of effective teachers, defined as a teacher's contribution to student learning (or value-added). The existing research shows an unequal distribution of teacher effectiveness based on factors such as teaching experience, teacher test scores, and certification in the subject area being taught (Presley et al. 2005; Lankford et al. 2002; Education Trust 2008; Clotfelter et al. 2006; Carroll et al. 2000). Although these studies use teacher characteristics as a proxy for teacher quality, the link between teacher characteristics and teacher effectiveness has not been well established (Rivkin et al. 2005; Gordon et al. 2006; Rockoff et al. 2008; Buddin and Zamarro 2008). Emerging evidence suggests that highly effective teachers are less likely to teach in schools with disadvantaged students, but the research base is small. The existing evidence is derived from Tennessee (Sanders and Rivers 1996; Tennessee Department of Education 2007) and the Dallas School District (Jordan et al. 1997).

As more states and school districts implement policies to address the inequitable distribution of effective teachers, there is a greater need to understand the scope of the problem and promising strategies to address it. In light of this need, IES has commissioned a study to measure the distribution of teacher effectiveness in a diverse set of 30 districts and assess the relationship between that distribution and the policies and programs in each district. Results of the study will provide educators, policymakers, and researchers with critical information on how teacher effectiveness is distributed and the role of policies to promote an equitable distribution.

b. Research Questions

The study's primary research questions are:

- What is the distribution of teacher effectiveness across schools within each district?
- What policies are districts using to address the inequitable distribution of teacher effectiveness?
- What is the relationship between district policies or strategies and the distribution of teacher effectiveness?

c. Study Overview

In this study, we will examine the distribution of teacher effectiveness in a diverse set of up to 30 school districts, document changes in the distribution over time, and analyze the relationship between the distribution and district policies designed to promote an equitable distribution.

We will purposefully select a diverse set of 30 districts to participate in the study. The primary criterion for study participation will be the availability of data to conduct a value-added analysis. While the number of states and districts with the data capacity needed for a value-added analysis is increasing, a statistical sampling method is not possible given the limited number of districts that currently have the data needed to estimate value-added measures. As a result, the sample of 30 districts selected for this study will not be nationally representative, and are likely to be unique because of the data capacity requirement. Additional criteria used to identify districts will include district size, geographic diversity, variation in equitable distribution policies, and socioeconomic diversity.

We will use a value-added model to estimate teacher effectiveness in each district. The value-added analysis will be conducted in two baseline years (2008-09 and 2009-10) and three subsequent years (2010-11 through 2012-13). The value-added measures will be used to measure the gap in effectiveness between the teachers of disadvantaged and non-disadvantaged students for each year, with student disadvantage measured by eligibility for FRL. We will calculate the Average Teacher Effectiveness Gap (ATEG), which is the difference in the average value added of the teachers of disadvantaged (i.e. eligible for FRL) and non-disadvantaged (ineligible for FRL) students, as well as the Average School Effectiveness Gap (ASEG), the difference in average value added of schools attended by disadvantaged and non-disadvantaged students. These measures represent the amount by which the teacher or school quality experienced by non-disadvantaged students differs from the teacher or school quality experienced by disadvantaged students.

The ATEG accounts for the distribution of teacher effectiveness both between and within schools, while the ASEG measures only the between-school component of this distribution. The ASEG relies only on school-student links rather than teacher-student links, and so can be computed even if there are incomplete data on teacher-student links. By comparing the value of the ATEG and ASEG, we can measure the extent to which the distribution of teacher quality is driven by differences in the value added of teachers across schools or within schools. The AEG is a flexible metric that can be used with a variety of measures of student disadvantage and can be used to trace the distribution of teacher quality for a cohort of students through multiple school years.

Interviews with district staff will provide information on district policies designed to address the distribution of teacher effectiveness for two baseline years and three subsequent years. The interviews will focus on how districts recruit, hire, evaluate, develop, and compensate them, and handle their transfers from school to school. We will document the timing of when districts implement these policies and quantify the key dimensions of district policies.

In this study, we will also analyze the role of district policy in the distribution of effective teachers. We will analyze variations in policies and distributional outcomes over three years both within and across districts. This correlational analysis is not designed to provide information about the effectiveness of these policies, but rather to offer new evidence about the relationship between strategies implemented by districts and the distribution of teacher effectiveness. As a complement to this analysis, and to help provide a more detailed explanation of the observed trends, we will also conduct a within-district analysis to examine the factors related to teacher mobility. For example, we will estimate the effect of average student poverty status within a school on a teacher's likelihood of remaining at the school.

d. Recruitment of Districts

To obtain a purposeful sample of 30 districts, we will begin by contacting 100 districts. Mathematica staff will contact the identified districts to gauge their interest in the study, confirm their data capacity for participation, and request their participation. We will begin the recruitment effort by mailing districts an introductory package, which will include the following two documents:

- **Notification letter.** The one-page notification letter on ED letterhead and signed by the contracting officer's representative describes the importance of studying the distribution of teacher effectiveness, provides an overview of the study design, summarizes the benefits of participating, and notes that a study team member will follow up by telephone to discuss the study in more detail (Appendix A).
- **Study summary.** The two-page summary describes the purpose of the study and the benefits of participation, identifies the study team, and provides contact information for the project director and the ED project officer. It also discusses the activities required of participating districts and schools (Appendix E).

We will send the notification letter and study summary to each district's superintendent and director of human resources via FedEx to highlight the importance of the documents. A Mathematica researcher will follow up with the director of human resources within two days of the delivery date to begin discussing the study. We will schedule in-person meetings or conference

calls with key stakeholders in the district to describe the study, explain the benefits of participation, confirm the availability of data needed for the study, discuss confidentiality procedures, and secure participation (Appendix B).

e. Data Collection Plan and Study Timeline

The study consists of two data collection efforts described below:

- District administrative data collection.** Mathematica will collect data from districts to conduct a value-added analysis and track teacher assignments and mobility. We will collect standardized test scores, student enrollment data with student-teacher links, and student demographic characteristics such as special education status and other factors that help explain test scores (Appendix C). The teacher personnel data include information on teachers' school assignments, movement within and out of the district each year, background characteristics, and teacher performance measures (Appendix C). Although we prefer to receive the data in electronic format, we will use data in whatever format is most convenient for each district. In the first round of data collection conducted in summer 2011, we will collect data for the past three school years (2007-08 through 2009-10). In the next three rounds of data collection, we will gather data for the next three school years (2010-11 through 2012-13). These three data collection rounds will begin in December 2011, 2012, and 2013 respectively.
- District staff interviews.** We will conduct telephone interviews with district staff who are knowledgeable about district policies designed to promote the equitable distribution of teacher quality. In the interviews, we will gather information on district policies related to the recruitment, hiring, transfer, evaluation, and compensation of teachers, as well as policies that affect school working conditions (Appendix D). Since it is unlikely that one person in the district will have sufficiently detailed information about each area, we will conduct interviews with three staff in each district. We will interview one staff person familiar with staffing (including recruitment, hiring, and transfer), another who is knowledgeable about compensation, and a third who is involved with school turnaround efforts. In the first round of interviews in summer 2011, we will gather data on district policies in the 2008-09 through 2010-11 school years. Interviews in spring 2012 will focus on district policies in the 2011-12 school year and those held in spring 2013 will concentrate on policies in the 2012-13 school year.

Table 1. Data Collection Plan

Activity	Summer 2011	Winter 2011	Spring 2012	Winter 2012	Spring 2013	Winter 2013
Collect student data and teacher personnel data from districts						
Data for 2007-08 through 2009-10 school years	X					
Data for 2010-11 school year		X				
Data for 2011-12 school year				X		

Data for 2012-13 school year				X
Conduct interviews with district staff				
Data on policies for 2008-09 through 2010-11 school years	X			
Data on policies for 2011-12 school year		X		
Data on policies for 2012-13 school year			X	

2. Purposes and Uses of Data

Mathematica and the Urban Institute will collect and analyze the data for the Study of the Distribution of Teacher Effectiveness under contract number ED-IES-10-C-0065.

The primary purpose of the study is to document the distribution of teacher effectiveness and changes in the distribution over time, and to analyze the role of district policies in distributional outcomes. This information will provide new evidence on the distribution of teacher effectiveness and inform the efforts of states and districts to track the distribution of teacher effectiveness and promote an equitable distribution. Through an analysis of data across school districts, the study will offer insights into the relationship between equitable distribution policies and the distribution of teacher effectiveness. The study will also provide information on how teacher mobility patterns affect the distribution of teacher effectiveness.

The data to be collected will be obtained from student records, teacher personnel data, and district staff interviews:

- **Student records data.** We will use existing state and district test score data, enrollment data with student-teacher links, and student background data to conduct a value-added analysis of teachers in the study districts. The teacher effect estimates from the value-added analysis will be used to measure the distribution of teacher effectiveness in each district and changes in the distribution over time.
- **Teacher personnel data.** These data will be used to analyze the role of teacher mobility and attrition for the distribution of teacher effectiveness. We will also use the data to explore the association between teachers' value-added scores and their decisions to transfer schools or exit the district.
- **District staff interviews.** We will use the interview data to examine the relationship between district policies and changes in the distribution of teacher effectiveness over time. These data will

also be used to document the types of policies that districts use to promote an equitable distribution of teachers.

3. Use of Technology to Reduce Burden

We will examine district websites and the Data Quality Campaign website to determine the nature of data available at the district and state levels. We will also request website links from districts to collect additional details about the policies and strategies not uncovered during preliminary website searches. Where feasible, we will gather information from existing databases, including student achievement test scores and demographic information. This information will be obtained in the form of computer files provided by school districts. We will provide clear instructions on the data requested and methods of transmitting the data securely. If it is too burdensome or not possible for a district to provide this information electronically, we will ask the district to provide hard copies of the relevant information, which will be coded by the study team. A program analyst will assist district staff in transferring data, as needed.

To track progress of data collection and minimize multiple contacts with districts, we will develop custom SharePoint lists and summary views to monitor both completion of the semi-structured interviews and acquisition and processing of student and teacher records from the districts.

Data from the semi-structured interviews cannot, however, be collected through such methods as web surveys or computer-assisted telephone interviews. The proposed telephone interviews will be necessary to allow in-depth, conversational exchanges with respondents, and to obtain answers to both open-ended and detailed questions. Prior to conducting the interviews, we will obtain information on district policies from the Teacher Rules, Roles, and Rights database developed by the National Council on Teacher Quality.

4. Efforts to Avoid Duplication

No national study has been conducted or is under way to address the research questions presented in this study. The Policy and Program Studies Service (PPSS) of ED is conducting a study of state and local teacher quality policies. The PPSS study will focus on gathering information about the design, development, and use of teacher quality measures. PPSS will not conduct interviews or surveys in the districts participating in the DTE study in order to reduce the burden on these districts. IES and PPSS will work together to ensure that the two studies do not duplicate efforts. Four studies have examined the distribution of teacher effectiveness, two of which researched the prevalence of high value-added teachers in Tennessee (Sanders and Rivers 1996; Tennessee Department of Education 2007). The third study examined the distribution in Dallas (Jordan et al. 1997). These studies found that schools with a high proportion of poverty and minority

students had a lower percentage of highest-performing teachers and a greater percentage of lowest-performing teachers as measured by value-added. A forthcoming study will examine the distribution of teacher effectiveness in ten districts, but will be limited to one point in time (Glazerman and Max, forthcoming). To date, this type of analysis has not been conducted on a larger or national scale, and there is no alternative source for the information to be collected.

5. Methods to Minimize Burden on Small Entities

We will minimize burden for school districts in the study by using existing data whenever possible and by requesting only the minimum data required to meet study objectives. Burden on districts will be further minimized through the careful specification of information needs. Data request documents will include the list of data items needed for the study (Appendix C).

We have developed an efficient interview protocol that focuses on the data of most interest. We will also speak with relatively few respondents in person—during in-person visits to recruit districts. The sample size and data requirements were determined by careful consideration of the information needed to meet study objectives and will be reviewed by the study’s technical working group (TWG).

6. Consequences of Not Collecting Data

The data collection activities described in this submission are necessary for ED to document the distribution of teacher effectiveness within districts over time and any changes in that distribution associated with district policies to promote an equitable distribution of teachers. The study represents a significant step in examining how low- and high-performing teachers are distributed on a larger and more national scale. Extending the data collection over time will allow us to implement the longitudinal data model relating district policies to the distribution of teacher effectiveness over the five-year period of the study. Without the data collected in this study, ED will not attain an understanding of the distribution of teacher effectiveness and any changes related to district policies.

7. Special Circumstances

There are no special circumstances associated with this data collection.

8. Federal Register Announcement and Consultation

a. Federal Register Announcement

The 60-day notice to solicit public comments was published in the *Federal Register*, on January 19, 2011 (Volume 76). The study did not receive any public comments. The 30-day notice to solicit public comments was

published in the *Federal Register*, on April 13, 2011 (Volume 76). One set of public comments were received on 5/13/11 from the National Education Association. A copy of the responses are included in a separate file.

b. Consultations Outside the Agency

The study team will work with IES to identify experts in teacher quality, value-added analysis, and evaluation methodology to become members of the TWG. Once they have been determined, we will seek their input on the study's design.

c. Unresolved Issues

There are no unresolved issues.

9. Payments or Gifts

We do not plan to give gifts to districts for completing the interview or providing other study data.

10. Assurances of Confidentiality

The data collection efforts that are the focus of this clearance package will be conducted in accordance with all relevant regulations and requirements, including the following:

- The Privacy Act of 1974, P.L. 93-579 (5 U.S.C. 552a).
- The Family Educational and Rights and Privacy Act (FERPA) (20 U.S.C. 1232g; 34 CFR Part 99).
- The Education Sciences Reform Act, P.L. 107-279 (20 U.S.C. 9573).

Mathematica and the Urban Institute will protect the confidentiality of information for the study and will use it for research purposes only. The project director will ensure that information about study members remains confidential. All data will be kept in secured locations. All members of the study team having access to the data will be trained and certified on the importance of confidentiality and data security. When reporting the results, we will present data in aggregate form only so that individuals and institutions will not be identified. We will also include the following statement in the requests for data:

- Responses to the data collection activities will be used for research purposes only. The reports prepared for the study will summarize findings across the sample and will not associate responses with a specific school or individual. We will not provide information that identifies you or your district to anyone outside the study team, except as required by law.

- The contractor follows the confidentiality and data protection requirements of the Institute of Education Sciences (The Education Sciences Reform Act of 2002, Title I, Part E, Section 183). The contractor will protect the confidentiality of all information collected for the study and will use it for research purposes only. No information that identifies any study participant will be released. Information on respondents will be linked to their institution but not to any individually identifiable information. No individually identifiable information will be maintained by the study team. All institution-level identifiable information will be kept in secured locations and identifiers will be destroyed as soon as they are no longer required.

The following safeguards, which are routinely employed by Mathematica to carry out confidentiality assurances, will be applied consistently during the study:

- All employees sign a confidentiality pledge (Appendix F), which describes both the importance of and the employee's obligation to discretion.
- Access to hard copy documents is strictly limited. Documents are stored in locked files and cabinets, and discarded materials are shredded.
- Access to computer data files is protected by secure user names and passwords, which are available to specific users only.
- Especially sensitive data is encrypted and stored on removable storage devices that are kept physically secure when not in use.

The plan for maintaining confidentiality includes ensuring that all personnel with access to study data sign confidentiality agreements and provide notarized nondisclosure affidavits. Also included in the plan is personnel training regarding (1) the meaning of confidentiality, particularly as it relates to handling requests for information and providing assurance to respondents about the protection of their responses; (2) controlled and protected access to computer files under the control of a single database manager; (3) built-in safeguards concerning status monitoring and receipt control systems; and (4) a secured and operator-controlled, in-house computing facility.

11. Justification for Sensitive Questions

The interviews, which will gather information about the design and implementation of school district policies, contain no questions of a sensitive nature. Test scores and some demographic information about the students may be sensitive. Test score data is essential for this study because student achievement is a key variable in our estimation of value-added. Demographic information is important to control for differences in student

characteristics across classrooms that may have arisen by chance, and thus is also a necessary input in the value-added analysis. For the teacher personnel data, we will gather background information such as date of birth, years of experience, certification status, and undergraduate education to control for factors that potentially affect mobility beyond a teacher's value-added score.

12. Estimates of Burden Hours

Table 2 shows the estimated burden hours for district staff who will participate in data collection. These estimates are based on our experience collecting such data from district staff for similar studies.

The number of respondents targeted for data collection is provided in the column *Number of Targeted Respondents* in Table 2. There are three data requests (phone interview, student records data collection rounds 2-4, and teacher personnel data collection rounds 2-4) that will occur three times. All other data requests occur once (recruitment by phone, recruitment by site visit, student records data collection round 1, and teacher personnel data collection round 1). Therefore, the total number of respondents is 640 (the sum of the number of respondents for the data collections that occur once plus the sum of the number of respondents for data collections that occur three times multiplied by three). The number of annual responses is 213 (640 divided by three years). Lastly, the total burden hours for data collection are provided in the last column of the table (4,130). The number of annual burden hours is 1,377 (4,130 divided by three years).

Table 2. Estimated Response Time for Data Collection

Respondent/ Data Request	Number of Respondents	Expected Response Rate (%)	Total Number of Responses	Unit Response Time (Hours)	Total Response Time (Hours/Year)	Total Burden Time (Hours)
District staff						
Recruitment by phone (once)	100	100	100	2	200	200
Recruitment by site visit ^a (once)	30	100	30	2	60	60
Phone interview (three staff at each district; 3 rounds)	90	100	270	1	90	270
Student records data, Round 1 ^b (once)	30	100	30	20	600	600
Student records data, Rounds 2-4 ^b (three times)	30	100	90	16	480	1,440
Teacher personnel data, Round 1 (once) ^c	30	100	30	16	480	480

Teacher personnel data, Rounds 2-4 (three times) ^c	30	100	90	12	360	1,080
Total	340	640	640	12	2,270	4,130

^aOf the 100 districts initially contacted for recruitment by phone, we expect to conduct site visits to 30 of those districts.

^bWe assume that the collection of student records data will take more time in the first year (Round 1) than in subsequent years (Rounds 2-4). We estimate that providing the student records data will require 20 hours of time in the first year and 16 hours in later years.

^cWe assume that the collection of teacher personnel data will take more time in the first year (Round 1) than in subsequent years (Rounds 2-4). We estimate providing the personnel data will require 16 hours in the first year and 12 hours in later years.

13. Estimates of Cost Burden to Respondents

There are no direct costs for respondents.

14. Annualized Cost to the Federal Government

The estimated annual cost of the study to the federal government is \$1,590,379. The estimated total cost of the five-year contract is \$7,951,897.

15. Reasons for Program Changes or Adjustments

This is a new data collection.

16. Plans for Tabulation and Publication of Results

Our tabulation plans for the study include the following analyses:

- Describing district policies to address the distribution of teacher effectiveness.
- Measuring teacher effectiveness in the classroom (using value-added analysis).
- Measuring the distribution of teacher effectiveness.
- Analyzing the distribution of teacher effectiveness within and among districts.

a. Tabulation Plans

i. Describing District Policies to Address the Distribution of Teacher Effectiveness

We will construct a set of policy variables to analyze the relationship between district policies and the distribution of teacher effectiveness. During interviews, we will gather information on 16 policies and practices within the policy categories described in Table 3. The interviews will be conducted at multiple points in time to capture new policies and policy changes

implemented during the study period (2008-09 through 2012-13 school years).

We will consider three approaches to creating policy variables for the analysis. One approach is to create a single variable for each of the 16 policies listed in Table 3. The main drawback of this approach is that because there are only 30 districts in the analysis, this would leave few degrees of freedom in a statistical analysis. Another option is to define variables based on the seven policy types shown in Table 3. These variables would provide a broader perspective on whether a district implemented each type of policy, and would reduce the number of policy variables to seven. For each district, we will sum the number of policies implemented within each policy type to indicate the *extent* of policy implementation within each type. For individual policies coded as categorical variables based on the extent of implementation, we will create decision rules to determine the level of implementation for the overall policy type.¹

Table 3. District Policies

Policy Type	Policies and Practices
Teacher recruitment and hiring	Recruitment activities that target high need schools Programs that recruit and hire teachers for high need schools (for example, TFA, Teaching Fellows, and so forth) Training and resources to improve recruitment and hiring by high need schools Early hiring timeline for high need schools
Teacher transfer	School choice in hiring voluntary transfers School choice in hiring involuntary transfers
Teacher compensation	Additional pay for teaching in high need schools Additional pay based on performance
Teacher layoffs	Teacher performance as a criterion in teacher layoffs Use of teacher performance as a factor in calling back and placing laid off teachers
Teacher development	Professional development activities targeted to teachers in high need schools Comprehensive induction programs that provide mentoring or support for new teachers
Teacher tenure	Teacher performance as a criterion for tenure decisions
School improvement policies	School turnaround activities (that is, closure, reconstitution, external management) Initiatives to improve working conditions in high need schools Initiatives to improve principal quality in high need schools

¹ For the extent of policy implementation variable, we will code each of the 16 policies as whether or not the district has such a policy (or has the policy at a particular level). The seven policy category variables will then be constructed as the sum of the number of policies within that category for a district.

Rather than grouping the individual policies based on the seven policy types, we can also group the individual policies based on which policies are typically implemented together by districts. An exploratory factor analysis (EFA) would be used to determine which policies are highly correlated (that is, are implemented in similar combinations across districts) and could be grouped together in a policy package. For example, based on the EFA, we might determine that districts tend to enact targeted recruitment efforts, early hiring timelines, and additional pay for teaching in high need schools as a package of policies. The EFA is a data-driven procedure to identify the packages of policies most commonly used (or not used) together by the districts in the study, and to create variables based on these policy packages. We will explore using the results of the EFA to create a continuous scale that indicates a district's overall prevalence of policy implementation of a particular package of policies (that is, the scale represents the weighted average of the policy indicator variables in a particular package). A disadvantage of EFA is that it may yield cluster groups of policies for which there is no obvious connection, making the resulting summary variables more difficult to interpret. An alternative approach to combining the policy variables will be to use a more informal method for grouping those that share some common conceptual or theoretical characteristic.

ii. Measuring Teacher Effectiveness

To measure teachers' value added scores, we rely on a regression model that controls for a series of baseline student characteristics that could be related to academic achievement, which might otherwise be confounded with the assignment of students to teachers. Specifically, we assume that a student's posttest score depends on prior achievement, background characteristics, their teachers, and additional unmeasured factors that are unrelated to teaching assignments. For each district, grade level, and subject (math or reading), the regression equation can be expressed as:

$$(1) \quad Y_i = \lambda W_i + \eta X_i + \theta T_i + \varepsilon_i$$

where Y_i is the posttest score for student i , and W_i represents a vector of pretests, including, at minimum, the test score for that student in the same subject in the prior year. The pretest scores capture prior inputs into student achievement. Control variables for student background characteristics are included in X_i , and T_i represents a set of variables for the teachers in the sample. Finally, ε_i represents an error term.

In an educational setting in which students are taught by only one teacher for the entire school year, T_i would contain a series of binary variables that indicate the link between teachers and students. For example, if teacher j had sole instructional responsibility for student i , then the j th element of T_i would be one and the remaining elements would be zero.

Similarly, the average effect of teacher j on the achievement of his or her students, after controlling for the other variables in the above equation, is represented in this case by θ_j , the j th element of the coefficient vector θ .

In our main analyses, we will control for a set of student characteristics, X_i , that is common to all study districts. This approach ensures that any differences we document in the distribution of teacher effectiveness across districts are not a result of using a different statistical model. The common value-added model will include the following student characteristics, which are generally available from district administrative records:

- Math and reading scores from the prior school year (regardless of the posttest subject)
- Free or reduced-price lunch eligibility
- Limited English proficiency
- Special education status
- Gender
- Student race/ethnicity

We will also check the sensitivity of the results by estimating district-specific models that include additional control variables not available in all districts. For example, we may include additional student background characteristics available in a subset of districts, such as parental education levels, days in attendance in the prior year, or multiple years of pretest scores.

We are planning to base our analysis of the distribution of teacher effectiveness on the Average Teacher Effectiveness Gap (ATEG), the average difference between the value-added scores of the teachers of disadvantaged and non-disadvantaged students within district, grade, year, and subject. Additional details about the ATEG are provided in the next section. This measure will be analyzed separately for math and for reading. Within each subject, we will aggregate the grade-specific ATEG across grades within a district-year and then compare this aggregate ATEG across years within a district and across districts within a year.

Because we are interested in measuring the ATEG each year, we will estimate value added using only the data for that year. In some contexts, researchers use multiple years of data to estimate teacher value added to generate more precise estimates of the component of teacher effectiveness that remains constant (McCaffrey et al. 2009). For our analysis, however, multi-year estimates could be disadvantageous if they mask true changes in teacher effectiveness from year to year, as we seek to quantify changes over time. Further, the ATEG pools information from multiple teachers, which implies that the precision gains to using multiple years of data are expected to be far smaller than if we were considering individual teacher value-added

measures. Considering both the bias and the precision, we believe that using single-year measures will increase the overall accuracy of ATEG estimates for our distribution analysis.

Because value-added measures are designed to make comparisons only among teachers whose students took the same pretests and posttests, we will convert all scale scores to z-scores by subtracting the mean and dividing by the standard deviation of a reference group. We assume that the underlying distribution of student ability is constant across grades, districts, and time. As a result, normalizing student scores within each grade, district, year, and subject reference group will yield value-added measures that are always stated in terms of standard deviations of student ability, or effect size units. This implies that the ATEG will also be expressed in effect size units of student achievement, which allows us to make the desired comparisons across districts and time. We will conduct sensitivity analysis to check if our results are robust to alternate ways of aggregating grade-level ATEG measures across grades, including (1) using alternate reference groups for defining the population mean and standard deviation, and (2) expressing effect sizes in terms of the distribution of teacher quality instead of the distribution of student achievement.

iii. Measuring the Distribution of Teacher Effectiveness

The primary goal of the study is to document the distribution of teacher effectiveness in a diverse set of districts. The measure of the distribution of teacher effectiveness will be used both to describe the distribution in each district and the key outcome variable in the analysis of the relationship between the distribution and district policies. The goal of the distribution measure is to describe the extent to which disadvantaged and non-disadvantaged students have equal access to effective teaching, defined here as teachers with high value-added scores.

The Average Effectiveness Gap (AEG) is a summary measure of the distribution of teacher effectiveness between disadvantaged and non-disadvantaged students, as defined by FRL status. We define two variants of this measure, the Average Teacher Effectiveness Gap (ATEG) and Average School Effectiveness Gap (ASEG), that depend on the level at which we are measuring effectiveness. The ATEG accounts for the distribution of teacher quality both between and within schools, and so is generally preferred to the ASEG, which measures only the between-school component of this distribution. The ASEG, however, relies only on school-student links rather than teacher-student links, and so can be computed even if there are incomplete data on teacher-student links.

The ATEG is the average value added of the teachers of non-disadvantaged students, V_{nd}^t , minus the average teacher value added of teachers of disadvantaged students, V_d^t :

$$(2) \quad ATEG = V_{nd}^t - V_d^t$$

Teachers who have both types of students in their classrooms will count toward both V_{nd}^t and V_d^t in proportion to the number of disadvantaged and non-disadvantaged students they have.

Similarly, the ASEG is the average value added of the schools attended by non-disadvantaged students minus the average value added of the schools attended by disadvantaged students. Schools that have both types of students count toward the average value added for both types of students.

These measures represent the amount by which the teacher or school quality experienced by non-disadvantaged students exceeds (if the AEG is positive) or is less than (if AEG is negative) the teacher or school quality experienced by disadvantaged students. It is numerically equal to the coefficient on FRL in the following regression of value-added scores on FRL:

$$(3) \quad V_{ij}^t = \alpha + \delta FRL_{ij} + e_{ij}$$

where V_{ij}^t is the value added of teacher or school j of student i , regressed on FRL_{ij} , a binary variable that takes a value of one if the student is *not* FRL-eligible and zero if the student is eligible. The estimated coefficient δ measures the estimated mean difference in teacher or school effectiveness between non-disadvantaged and disadvantaged students in the district, with a positive δ indicating an inequitable distribution and a negative δ indicating a compensatory distribution. The AEG can be used as a credible measure of the distribution of teacher or school effectiveness regardless of whether the value-added model used to generate measures of teacher or school effectiveness includes FRL as a control variable.

Within- and Between-School Differences in the Distribution of Teachers. By comparing the value of the ATEG and ASEG, we can measure the extent to which the distribution of teacher quality is driven by differences in the value added of teachers across schools or within schools. This can be useful for diagnosing the source of an inequitable distribution of teachers, suggesting whether policymakers in a particular school district would be better off focusing on policies like hiring and retention reforms that equalize teacher quality across schools or are better off focusing on tracking policies that determine teacher-student matches within schools.

- If the ASEG is larger than the ATEG, that suggests that principals assign disadvantaged students to the higher value-added teachers within schools. The matching of students to teachers across schools is the source of any inequity in the distribution of teacher

effectiveness, with within-school assignment serving as a compensatory mechanism.

- If the ASEG and ATEG are equal, this suggests that any difference in teacher effectiveness (whether inequitable or compensatory) is across schools, and that teacher assignments within school, although not compensatory, are not responsible for the overall difference in the distribution of teacher effectiveness.
- If the ATEG is larger than the ASEG, that suggests that within-school sorting of students is exacerbating any inequity in the distribution of teacher effectiveness associated with between-school sorting. For example, if the ASEG were zero and the ATEG were positive, this would suggest that all of the inequity is due to within-school sorting of students to teachers.

The degree to which the ASEG and ATEG can differ depends on the amount of segregation of non-disadvantaged and disadvantaged students across schools. For example, if non-disadvantaged and disadvantaged students are completely segregated into different schools, there will be a negligible difference between the two measures, as there would be little opportunity for principals to differentially assign disadvantaged or non-disadvantaged students to different teachers within schools. This would also be the case if teacher quality did not vary within schools, even if disadvantaged and non-disadvantaged students were well integrated. To make the AEG more useful to policymakers seeking to target policies to redress inequities, we will present the ATEG alongside the ASEG, the percentage of students in the district who are FRL-eligible, the across-school and within-school variation in teacher quality, and a measure of student segregation across schools.

Extensions of the AEG. The AEG is a flexible metric that can be used with a variety of measures of student disadvantage and can be used to trace the distribution of teacher quality for a cohort of students through multiple school years.

One extension would be to replace FRL with an alternate measure of inequality. For example, by replacing FRL with race/ethnicity, we can use a similar analysis to measure black/white or Hispanic/white gaps in the distribution of effective teachers. By replacing a binary variable of student disadvantage with the student pretest score (ideally measured before a student enters a school), we can extend the AEG to a case where the measure of student inequality is a continuous rather than a discrete variable.

A second extension to measuring the AEG in a single year would be to consider the AEG as it affects students over multiple years. Because we will collect data that allow us to compute value-added in a school district over five years, we can trace the effectiveness of teachers of one cohort of students for every year between grades 4 and 8, and for other cohorts for multiple years within the range of grades 4 to 8. By measuring the AEG

annually over multiple years, we will be able to examine whether there is a cumulative gap that grows larger each year or whether inequitable distributions in some years are offset by compensatory distributions in others.

A third extension that would serve as a sensitivity test, is to calculate the AEG based on a district-specific measure of teacher performance rather than a common value-added model. While a common value added model provides a consistent performance measure for all participating districts, a district's policies may be based on a different type of measure, such as a classroom observation rubric, or based on a combination of measures, such as a classroom observation rubric and student achievement growth measure. We will request teacher performance measures as part of the teacher personnel data request, collecting teacher evaluation results or other evaluation tools that lead to a numeric rating of teachers, as well as performance measures that are used as the basis for teacher compensation, teacher tenure, or teacher layoff policies. When these data are available, we will test the sensitivity of the distribution results and the policy analysis described in the next section, to see if district policies have a stronger relationship with the distribution when measured with district-specific measures of teacher performance.

Finally, rather than focusing only on the average gap, one can plot a histogram of teacher effectiveness for non-disadvantaged students and the same for disadvantaged students. The effectiveness of teachers who teach both types of students will be represented in both histograms, weighted by the number of disadvantaged and non-disadvantaged students they teach. This will show the degree of overlap, and the degree to which average differences may be due to a greater likelihood of one group or another being assigned to teachers at the tails of the distribution. Sass et al. (2010) use this technique to show how the distribution of teacher value added compares across schools with 70 percent or more FRL students compared to 70 percent or fewer FRL students, and find a thicker tail of ineffective teachers in the higher-poverty schools. Our approach would be slightly different, as the histograms would represent teachers of disadvantaged students in all schools compared to teachers of non-disadvantaged students in all schools.

iv. Analyzing the Relationship Between Policies and the Distribution

The goal of this analysis is to provide an initial understanding of the relationship between district policies and the teacher effectiveness gap. While all of these proposed analyses are exploratory (that is, non-causal), they can suggest policies and practices that should be examined using more rigorous methods in the future.

To summarize the relationship between district policies and the distribution of teacher effectiveness in a given subject, we will create a series of tables and graphs that illustrate the ATEG and ASEG mean and

standard deviation across those districts with and without each of the policies identified. In this analysis, we will incorporate all five years of data for each district to compare the average difference in ATEG and ASEG for districts that ever implemented a policy or groups of policies and districts that never implemented the policy. The difference in ATEG and ASEG scores will provide an initial sense for how the distribution in these districts differs, and we will test the significance of these differences.

We will also examine changes in the ATEG and ASEG over time to see whether the gap is improving across districts during the course of the study. We will analyze how the change in ATEG or ASEG over the course of the five-year period correlates with policy implementation. Also, we will account for whether districts implement policies for the entire study period (potentially including years prior to the study period), a portion of the study period, or never during the study period, with a variable indicating the number of study years in which a district implemented the policy (ranging from zero to five). We will correlate the change in ATEG from 2008-09 to 2012-13 with the number of years of implementation to initially examine the relationship between policy implementation and the ATEG.

We will extend the summary statistics described above by estimating a model that relates district policies to the distribution of teacher effectiveness. This model will include control variables that are likely to be correlated with both the policy variables as well as the ATEG. To measure the relationships between policies and the distribution, we would estimate:

$$(4) \quad ATEG_{it} = \alpha + \beta \mathbf{P}_{it} + \gamma \mathbf{X}_i + \lambda_t + u_i + \varepsilon_{it}$$

where $ATEG_{it}$ represents the average teacher effectiveness gap for district i in year t ; \mathbf{P}_{it} is a vector of policy variables that represent the district's implementation of policies in year t ; \mathbf{X}_i is a vector of time invariant district characteristics; λ_t captures time-specific shocks in the ATEG across districts; u_i is a district random effect term to adjust the standard errors associated with having multiple years of data for each district; and ε_{it} is an error term with subscripts i and t representing districts and time, respectively. In an alternate specification, we will also consider including district fixed effects instead of time invariant district characteristics (see below). Because some policies are expected to affect the distribution in the year after they are implemented, we will define some policies based on their implementation in the prior year. For example, a layoff policy implemented in year $t-1$ would be expected to have an effect on the distribution in year t .²

² In addition, we will also consider whether a lag is expected between when a policy is implemented and when it affects the distribution. For example, teacher induction programs were found to have an impact in the third year after implementation, but not in the first two years.

The analysis will control for a set of district characteristics that are potentially related to the distribution.³ This includes the geographic region of the district, the size of the district, and the level of urbanicity. Because the distribution of teacher effectiveness may vary based on whether the teacher labor market is a single district or a group of neighboring districts, we will include a measure of the extent to which students in a metropolitan area are concentrated in a single district.⁴ To account for the distribution of parents' income, we include the median household income and variability of household income in the district or metropolitan statistical area.⁵ We will also capture district activity that might be related to the distribution, but unrelated to the policies included in the analysis (for example, enrollment trends that could lead to changes in the number of schools). Finally, we will consider other contextual factors described during district staff interviews that might be associated with the distribution of teacher effectiveness.

v. Modeling changes to the distribution of teacher effectiveness through teacher mobility.

Changes over time in the distribution of teacher effectiveness can occur for a number of reasons: hiring of new teachers, within-district transfers of teachers, teachers leaving the district, changes in the effectiveness of teachers, and demographic changes across schools. The purpose of this analysis is to better understand the role of teacher mobility in changes in the distribution of teacher effectiveness. We will examine the relationship between a teacher value added and the probability that a teacher continues teaching at the same school, transfers to another school within the district, or leaves the district. We are especially interested in the relationship between attrition, value added, and school characteristics (for example, whether high value-added teachers are more likely to transfer out of high poverty or low achieving schools). In addition, we will look at the characteristics of schools that teachers move to, investigating whether high value-added teachers might see relatively larger improvements in school characteristics associated with a move.

³ Our model supports the use of time-varying district characteristics, and in situations where we are able to utilize time-varying data, we will do so. However, many of these district variables will remain stable over time, and many of the data sources we will use to track these variables are not updated frequently enough to monitor changes over time.

⁴ We will use the Herfindahl Index, which is calculated by taking the proportion of students from each district in the MSA, squaring each value, and taking the sum of these squared values. It is highly correlated with the percentage of students in the largest district in the MSA. As the number of equal sized districts grows larger, the Herfindahl Index approaches zero. As the number of dominant districts claiming a larger share of the region's population grows smaller, the index approaches one.

⁵ We would parameterize the level of income as the logarithm of the median income and the variability of income as the standard deviation of the logarithm of income, the Gini coefficient of income, or the percentage in poverty.

b. Publication Plans

We will prepare two reports and two evaluation briefs presenting the results of the tabulations described above. The first report, with a projected release in 2012, will address the distribution of teacher effectiveness across schools within districts for the baseline years. The two annual evaluation briefs will provide descriptive information about study sites' strategies to promote an equitable distribution of teachers and information on the distribution of teacher effectiveness within study districts. The first brief is expected to be released in 2013, and the second in 2014. The second report, with a projected release in 2015, will examine the association between districts' strategies and the distribution of teacher effectiveness within districts and discuss patterns over time. Reports will be written in a style and format that are accessible to policymakers and practitioners and will comply fully with the standards set by the National Center for Education Statistics. These study reports will be available in hard copy and on the ED website.

17. Approval Not to Display the Expiration Date for OMB Approval

The study will display the OMB expiration date.

18. Exception to the Certification Statement

No exceptions are being sought.

REFERENCES

- Boyd, Donald, Pam Grossman, Hamilton Lankford, Susanna Loeb, and James Wyckoff. "Who Leaves? Teacher Attrition and Student Achievement." National Bureau of Economic Research, Inc., NBER Working Papers: 14022, 2008.
- Buddin, Richard, and Gema Zamarro. "Teacher Quality, Teacher Licensure Tests, and Student Achievement." Santa Monica, CA: RAND, May 2008.
- Carroll, Stephen, Robert Reichardt, and Cassandra Guarino. "The Distribution of Teachers Among California's Districts and Schools." Santa Monica, CA: RAND, October 2000.
- Clotfelter, Charles, Helen Ladd, Jacob Vigdor, and Justin Wheeler. "High-Poverty Schools and the Distribution of Teachers and Principals." Washington, DC: National Center for Analysis of Longitudinal Data in Education Research, December 2006.
- Education Trust. "Their FAIR Share: How Texas-Sized Gaps in Teacher Quality Shortchange Low-Income and Minority Students." Washington, DC: EdTrust, February 2008.
- Glazerman, Steven, and Jeffrey Max. "Mapping the Prevalence of High-Performing Teachers." Washington, DC: Mathematica Policy Research, forthcoming.
- Gordon, Robert, Thomas Kane, and Douglas Staiger. "Identifying Effective Teachers Using Performance on the Job." Washington, DC: Hamilton Project, Brookings Institution, April 2006.
- Jordan, Heather, Robert Mendro, and Dash Weerasinghe. "Teacher Effects on Longitudinal Student Achievement." Paper presented at the CREATE Annual Meeting, Indianapolis, IN, July 1997.
- Lankford, Hamilton, Susanna Loeb, and James Wyckoff. "Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis." *Educational Evaluation and Policy Analysis*, vol. 24, no. 1, spring 2002, pp. 37-62.
- McCaffrey, Dan, Tim Sass, J.R. Lockwood, and Kata Mihaly. "The Intertemporal Variability of Teacher Effect Estimates", *Education Finance and Policy*, vol. 4, no. 4, Fall 2009, pp. 572-606.
- Presley, Jennifer, Bradford White, and Yuqin Gong. "Examining the Distribution and Impact of Teacher Quality in Illinois." Edwardsville, IL: Illinois Education Research Council, 2005.

- Reardon, S.F., and G. Firebaugh. "Measures of Multigroup Segregation." *Sociological Methodology*, vol. 32, January 2002, pp. 33-67.
- Rivkin, Steven, Erick Hanushek, and John Kain. "Teachers, Schools, and Academic Achievement." *Econometrica*, vol. 73, no. 2, March 2005, pp. 417-458.
- Rockoff, Jonah, Brian Jacob, Thomas Kane, and Douglas Staiger. "Can You Recognize an Effective Teacher When You Recruit One?" Cambridge, MA: National Bureau of Economic Research, November 2008.
- Sakoda, James. A Generalized Index of Dissimilarity. *Demography*, vol. 18, no. 2, 1981, pp. 245-250.
- Sanders, William, and June Rivers. "Cumulative and Residual Effects of Teachers on Future Student Academic Achievement." Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center, November 1996.
- Sass, Tim, Jane Hannaway, Zeyu Xu, David Figlio, Li Feng. "Value Added of Teachers in High-Poverty Schools and Lower-Poverty Schools." Washington, DC: The Urban Institute, November 2010.
- Tennessee Department of Education. "Tennessee's Most Effective Teachers: Are They Assigned to the Schools That Need Them Most?" Nashville, TN: Tennessee Department of Education, 2007.

MATHEMATICA
Policy Research, Inc.

**[www.mathematica-
mpr.com](http://www.mathematica-mpr.com)**

Improving public well-being by conducting high-quality, objective research and surveys

Princeton, NJ ■ Ann Arbor, MI ■ Cambridge, MA ■ Chicago, IL ■ Oakland, CA ■ Washington, DC

Mathematica® is a registered trademark of Mathematica Policy Research