Study of the Distribution of Teacher Effectiveness

Part B

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 Distributioni of Teacher Effectiveness

Part B

June 9, 2011

MATHEMATICA
Policy Research, Inc.

CONTENTS

SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT PART B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

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) and the U.S. Department of Education (ED) and implemented by Mathematica 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 to 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. 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.

1. Respondent Universe and Sampling Methods

Ideally, we would obtain a nationally representative sample of 30 districts from the universe of school districts in the country. However, this is not feasible because participating districts must have the data needed to conduct a value-added analysis. A value-added model requires data with student-teacher links that identify for each student the teacher responsible for teaching math and reading, as well as unique student and teacher identifiers that remain the same over time. While the number of states and districts with this data capacity is increasing, a statistical sampling method is not possible given the limited number of districts with these data. In addition, conducting value-added analyses, measuring the distribution of effective teachers, and analyzing teacher mobility requires that participating districts have a sufficient number of students, teachers, and schools (see Section B below).

To identify 30 districts based on these criteria, we will begin with a list of the largest 100 districts from the Common Core of Data maintained by ED's National Center for Education Statistics (NCES). Since the largest districts are disproportionately located in the South and West regions of the country, we

will expand the list to include districts among the next 600 largest that were located in the North or Midwest and had at least 30 total schools. This will provide an initial list of 160 districts for recruitment.

Since variation across districts in the types of teacher distribution policies will facilitate the analysis of how district policies relate to changes in the teacher distribution, we will use information on the following types of policies:

- **Teacher compensation**. We will use the Teachers Rules, Roles, and Rights (TR3) database maintained by the National Council on Teacher Quality and information from the websites of the Center for Educator Compensation Reform and the National Center on Performance Incentives websites to identify policies and programs that (a) offer additional pay or a higher salary for teaching in a high-need school, (b) reward teachers based on their performance, or (c) pay teachers for taking on additional roles or responsibilities.

- **Teacher recruitment**. Information on district policies or practices designed to recruit a pool of teachers for high-need schools. This includes targeted recruitment activities for high-need schools as well as programs such as Teach for America, Teaching Fellows, or Teacher Residency Programs that recruit new teachers for these schools.

- **School autonomy in personnel decisions**. The TR3 database will be used to gather information on district policies that provide schools the authority to decide whether to hire a voluntary or involuntary transfer and that require the consent of the principal and teacher when filling a vacancy.

Districts often implement more than one type of equitable distribution policy, so we will group districts into the following four policy categories: (1) teacher compensation policies only; (2) teacher recruitment policies only, or recruitment policies with compensation policies; (3) autonomy policies with recruitment and/or compensation policies; and (4) none of these three policies. Among the 160 districts identified based on district size, 21 percent are in the first policy category, 18 percent are in the second, 17 percent are in the third, and 44 percent are in the fourth. However, this overstates the proportion of districts in the fourth category because it also captures districts for which information on district policies was not available.

We will then prioritize an initial group of 100 districts for recruitment using the following approach:

- **Maximize geographic diversity.** While a nationally representative sample is not possible, we will seek representation

from each of the four geographic regions that is proportional to the relative number of students in each of the four major geographic regionsin the United States: North, South, Midwest and West.

- **Prioritize data capacity.** We will prioritize districts that have the data to support the estimation of teacher value-added indicators, including student-teacher links and unique identifiers for students and teachers that remain consistent over time.
- **Seek variation in equitable distribution policies.** We will seek districts that vary in the types of equitable distribution policies they implement. To do this, we will target districts from each of the four policy categories described above to ensure variation in this key dimension for the analysis. The extent of variation in policies across participating districts will depend on the data capacity of districts in each policy category.
- **Prioritize districts with greater socioeconomic diversity and geographic concentration.** When there are multiple districts meeting the above criteria, we will prioritize districts with more socioeconomic diversity, at least 20 percent and less than 80 percent of students eligible for free or reduced-price lunch, and more geographic concentration, districts where students are concentrated in one larger district rather than spread across multiple smaller districts.

Recruitment efforts will focus on the initial group of about 100 districts, and we will have a backup pool of 60 additional districts as needed when districts refuse to participate or lack the necessary data capacity.

2. Procedures for the Collection of Information

To obtain the purposeful sample of 30 districts, Mathematica staff will contact the identified districts to gauge their interest in the study and request their participation. We will begin the recruitment effort by mailing districts introductory packages, 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.

We will collect the following data from the 30 districts participating in the study:

- **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 of 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, development, and compensation of teachers, as well as policies that affect school working conditions (Appendix D). Also, we will conduct interviews with three staff in each district who are knowledgeable in these areas. Since it is unlikely that one person in the district will have sufficiently detailed information about each area, we will interview one staff person familiar with staffing (that is, 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.

The sample size requirements for the study were developed by identifying the numbers of students, teachers, and schools necessary to answer the study's main research questions with a reasonable degree of precision. We used past research on similar topics to inform our sample size requirements. The sections below describe each analysis and the sample size requirements.

a. 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:

$$Y_i = \lambda' \mathbf{W}_i + \eta' \mathbf{X}_i + \theta' \mathbf{T}_i + \varepsilon_i \quad (1)$$

where Y_i is the posttest score for student i, and $\mathbf{W}_i$ 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 $\mathbf{X}_i$, and $\mathbf{T}_i$ represents a set of variables for the teachers in the sample. Finally, ε_i represents an error term.

Because we are interested in measuring the distribution 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, our distribution measure described below 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 the estimates for our distribution analysis.

Therefore, we require that school districts included in the study be able to initially supply student test scores for two years and teacher-student links for one year—the data required to estimate one year of value-added teacher estimates.

b. 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^t_{nd}, minus the average teacher value added of teachers of disadvantaged students, V^t_d:

(2) $$ATEG = V^t_{nd} - V^t_d$$

Teachers who have both types of students in their classrooms will count toward both V_{nd} and V_d 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) $V_{ij}^{t} = \alpha + \delta FRL_{ij} + \boldsymbol{e}_{ij}$,

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.

A reliable measure of the distribution of teacher effectiveness based on the ASEG requires a minimum of three schools per grade span. Therefore, we require that participating school districts have at least three elementary schools and three middle schools

c. Among-District Analysis

We will estimate a longitudinal model relating district policies to changes in the distribution of teacher effectiveness. 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. We will construct a set of policy variables to analyze the relationship between district policies and the distribution of teacher effectiveness. We will estimate 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) $$ATEG_{it} = \alpha + \boldsymbol{\beta}' \mathbf{P}_{it} + \boldsymbol{\gamma}' \mathbf{X}_i + \lambda_t + u_i + \varepsilon_{it},$$

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*.[1]

In order to identify statistically significant correlations between policies and the distribution of teacher effectiveness, we require at least 30 districts to provide the sample size necessary to conduct this analysis.

d. Within-District Analyses

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.Past research tracing the mobility of teachers with value-added measures of performance has relied on state databases (Goldhaber et al. 2007) or data from single, large school districts. (Hanushek et al. 2005 use data from a single anonymous Texas district; Boyd et al. 2007 use data from New York City.) In order to extend these types of analyses, we intend to follow the lead of these authors who used data from a large school district. Because our sample of districts will include 30 districts that are among the largest in the country, we are confident that we will have a sufficiently large sample for this analysis.

[1] 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.

3. Methods to Maximize Response Rates and Deal with Nonresponse

Recruiting for the study. We will rely on several strategies to secure school districts' participation in the study. The recruiters will be trained to present information clearly, address concerns, and respond to questions quickly and effectively. We will use ED letterhead for the notification letters to more readily capture districts' attention, and recruiters will indicate that they are calling on behalf of ED when they speak to representatives of the districts. Recruiters will also inform district representatives that the study meets all federal research guidelines, and was reviewed by both the Office of Management and Budget (OMB) and an independent institutional review board (IRB). The recruitment task leader will monitor recruiting issues daily to quickly resolve any obstacles to participation that might arise.

Data collection. Mathematica has developed multiple strategies to maximize response rates while minimizing burden on respondents. We have found that the following techniques contribute significantly to a high completion rate: establishing positive relationships with respondents and district staff; sending advance letters; and establishing efficient and flexible scheduling. To help alleviate districts' concerns about data privacy, all information request documents will include a statement on our adherence to confidentiality and data collection requirements (The Education Sciences Reform Act of 2002, Title I, Part E, Section 183).

District leaders are likely to have a strong interest in piloting a value-added model and using the resulting measures to understand the distribution of effective teachers. Recent federal investments through ARRA and other sources have shifted the focus from teacher qualifications to teacher effectiveness. Therefore, participating in this study, with its focus on the distribution of teacher effectiveness, will be more attractive to district leaders. We will begin our recruitment with 50 districts with the understanding that some will decline our initial recruitment efforts or not have the data capacity to participate. We will then collect data from the 30 districts that agree to participate in the study and cooperate with data collection. To further solidify the cooperation of district staff, we will adhere to any data collection requirements that districts may have, such as preparing research applications and seeking IRB approvals. Once we identify the 30 districts needed for the study, we anticipate a 100 percent response rate for district administrative records and staff telephone interviews.

4. Tests of Procedures or Methods to Be Undertaken

When possible, data collection instruments for the study will be modeled on protocols and data requests that have been used successfully with previous studies, such as the School District Questionnaire for ED's School and Staffing Survey (OMB Control Number 1850-0598) and the data request form for the Moving High Performing Teachers to Low Performing Schools

Evaluation (OMB Control Number 1850-0861). The pilot will assess the content, clarity, and wording of individual questions; respondent burden time; and the use of probes. The data collection instruments and results of the pre-tests will be included in the final OMB package.

In addition to pilot testing specific questions, we will pre-test the interview protocol with a total of six staff in two school districts to identify problems respondents might have in providing the requested information. We will pre-test the data request form with a staff person from the data office within each pilot district and the interview protocol with three staff in the two pilot districts.

5. Individuals Consulted on Statistical Aspects of the Design and on Collecting and/or Analyzing Data

The following individuals were consulted on the statistical aspects of the study design and on the data collection and analysis:

Name	Title	Telephone Number
Philip Gleason	Senior Fellow, Mathematica	315-781-8495
Steve Glazerman	Senior Fellow, Mathematica	202-484-4834
Eric Isenberg	Researcher, Mathematica	202-554-7540
Jeffrey Max	Researcher, Mathematica	202-484-4236
Mary Grider	Systems Analyst, Mathematica	202-484-4820

REFERENCES

Boyd, Donald, Pam Grossman, Hamilton Lankford, Susannah Loeb., and James Wyckoff. “Who Leaves? Teacher Attrition and Student Achievement.” Washington, DC: Urban Institute, National Center for Analysis of Longitudinal Data in Education Research, May 2008.

Goldhaber, Dan, Gross, Bethany Gross, and Dan Player. “Are Public Schools Really Losing Their “Best”? Assessing The Career Transitions Of Teachers And Their Implications For The Quality Of The Teacher Workforce.” Working Paper 12. Washington, DC: Urban Institute, National Center for Analysis of Longitudinal Data in Education Research, October 2007.

Hanushek, Eric, John Kain, Daniel O’Brien, and Steve Rivkin. “The Market for Teacher Quality.” Working Paper No. 11154. Washington, DC: National Bureau of Economic Research, February 2005.

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.

Reardon, S.F., and G. Firebaugh. Measures of Multigroup Segregation. Sociological Methodology, vol. 32, January 2002, pp. 33-67.

Sakoda, James. A Generalized Index of Dissimilarity. Demography, vol. 18, no. 2, 1981, pp. 245-250.

Sass, Tim, Jane Hannaway, Zeyu Xu, David Figlio, Li FengValue Added of Teachers in High-Poverty Schools and Lower-Poverty Schools. Washington, DC: The Urban Institute, November 2010.

Schochet, Peter and Hanley Chiang. "Error Rates in Measuring Teacher and School Performance Based on Student Test Score Gains." Washington, DC: Mathematica Policy Research, July 2010.

Theil, Henri. Statistical decomposition analysis. Amsterdam: North-Holland Publishing Company, 1972.