



4-1-2008

Educational Considerations, vol. 35(2) Full Issue

Chad Litz
Kansas State University

Follow this and additional works at: <https://newprairiepress.org/edconsiderations>



Part of the [Higher Education Commons](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-Share Alike 4.0 License](#).

Recommended Citation

Litz, Chad (2008) "Educational Considerations, vol. 35(2) Full Issue," *Educational Considerations*: Vol. 35: No. 2. <https://doi.org/10.4148/0146-9282.1191>

This Full Issue is brought to you for free and open access by New Prairie Press. It has been accepted for inclusion in Educational Considerations by an authorized administrator of New Prairie Press. For more information, please contact cads@k-state.edu.

ISSN No.
0146-9283

**Spring
2008**

educational considerations

published at kansas state university college of education

WE'RE ONLINE!

***Your subscription to
Educational Considerations
is now available online.***

YOU HAVE CHOICES!!!

***#1: Continue to receive paper copy
at your regular subscription price.***

***#2: Receive BOTH paper and electronic copy
at your regular subscription price.***

***#3: Receive ONLY ELECTRONIC copy at
80% of regular subscription price.***

***To get your online password, e-mail us at thomsond@ksu.edu.
Tell us in your e-mail whether you want paper AND electronic
copy, or whether you want only electronic copy.***

***VISIT US AT:
<http://coe.ksu.edu/EdConsiderations/>***

educational considerations

Vol. XXXV, Number 2, Spring 2008

Available online at:
<http://coe.ksu.edu/EdConsiderations/>

BOARD OF EDITORS

Chad Litz, Chair Emeritus

Kansas State University

David C. Thompson

Kansas State University

R. Craig Wood

University of Florida

G. Daniel Harden

Washburn University

EXECUTIVE EDITOR

Faith E. Crampton

University of Wisconsin-Milwaukee

Mary L. Hammel

Assistant to the Editor, Kansas State University

EDITORIAL ADVISORY BOARD

Donald J. D'Elia

State University of New York-New Paltz

Patrick B. Forsyth

Oklahoma State University

William Fowler

George Mason University

Janis M. Hagey

National Education Association

William Hartman

Pennsylvania State University

Marilyn Hirth

Purdue University

Richard King

University of South Florida

Martha McCarthy

Indiana University

Mary McKeown-Moak

MGT of America, Inc.

F. Howard Nelson

American Federation of Teachers

Allan Odden

University of Wisconsin-Madison

Margaret L. Plecki

University of Washington

Catherine Sielke

University of Georgia

William E. Sparkman

University of Nevada-Reno

Julie Underwood

University of Wisconsin-Madison

Deborah A. Verstegen

University of Nevada-Reno

James G. Ward

University of Illinois-Champaign-Urbana

TABLE OF CONTENTS

Optimizing Educational Resources: A Paradigm for the Pursuit of Educational Productivity	3
James L. Phelps	
Teacher Compensation and School Quality: New Findings from National and International Data	19
Zhijuan Zhang, Deborah A. Verstegen, and Hoe Ryoung Kim	
Explaining the Relationship Between Resources and Student Achievement: A Methodological Comparison of Production Functions and Canonical Analysis	29
Robert C. Knoeppel and James S. Rinehart	
Funding Michigan K-12 Adequacy Without Rewarding Inefficiency	41
James J. Walters	
Commentary	
Unacceptable but Tolerated Behavior	50
Anne L. Jefferson	

Educational Considerations invites subscribers for only \$13.00 annually. **Educational Considerations** is published and funded by the College of Education at Kansas State University. Write to EDITOR, Bluemont Hall, Kansas State University, Manhattan, KS 66506 or call (785) 532-5543.

PUBLICATION INFORMATION

Educational Considerations is a peer-reviewed journal published at the College of Education, Kansas State University. **Educational Considerations** and Kansas State University do not accept responsibility for the views expressed in articles, reviews, and other contributions appearing in this publication. In keeping with the professional educational concept that responsible free expression can promote learning and encourage awareness of truth, contributors are invited to submit conclusions and opinions concerned with varying points of view in and about education.

Educational Considerations is published two times yearly. Editorial offices are located at the College of Education, Bluemont Hall, 1100 Mid-Campus Drive, Kansas State University, Manhattan, KS 66506-5301. Correspondence regarding manuscripts should be directed to the Executive Editor

at fec@uwm.edu. No remuneration is offered for accepted articles or other materials submitted.

By submitting to **Educational Considerations**, the author guarantees that the manuscript has not been previously published. The University of Chicago's *Manual of Style* is the editorial style required. Footnotes should be listed at the end of the manuscripts. Tables, graphs, and figures should be placed in separate files. Camera ready hard copy should be available upon request. An abstract of 150 words must accompany the manuscript. **Manuscripts should be submitted electronically to Faith Crampton at fec@uwm.edu as an e-mail attachment.** Complete name, address and telephone number should be included in the body of the e-mail. Photographs, drawings, cartoons, and other illustrations are welcome. Authors should be prepared to provide copies of permission to quote copyrighted materials. Queries concerning proposed articles or reviews are welcome. The

editors reserve the right to make grammatical corrections and minor changes in article texts to improve clarity. Address questions regarding specific styles to the Editor.

Subscription to **Educational Considerations** is \$13.00 per year, with single copies \$10.00 each. Correspondence about subscriptions should be addressed to the Business Manager, c/o The Editor, **Educational Considerations**, College of Education, Kansas State University, Manhattan, KS 66506-5301. Checks for subscriptions should be made out to **Educational Considerations**.

Printed in the United States of America.

*Educational Considerations Design/Layout by
Mary Hammel, Kansas State University*

Optimizing Educational Resources: A Paradigm for the Pursuit of Educational Productivity

James L. Phelps

The advantage and perhaps the major motivation for using “seat-of-the-pants” decision making is that it obscures the assumptions made in arriving at a decision. If no one knows the assumptions upon which you based your decisions, then even though they may be uneasy with the decision they will have a difficult time criticizing your assumptions or decisions. (Schrage, p. 305)¹

The never-ending organizational challenge is to allocate available resources to best achieve its goals. Out of this fundamental question several models have evolved. One is a conceptual model—a way to think about how organizations operate. A second is a statistical model estimating the magnitude of relationships among goals and elements of the organization. This article presents a third model, an optimization model building upon the other two in order to analyze various policy options by simulating “what if” situations arising in organizations. These three models are complementary rather than competing.

Optimization Modeling

What Is a Model?

Over time, scientific endeavors have increasingly relied on models combining fact (observations), theory (assumptions), laws (usually mathematical), and methodology (procedures) into a system describing phenomena behavior. Models evolve as anomalies, are identified in older models, and are replaced with different facts, theories, laws, and methodologies describing the behavior of the phenomenon in question more comprehensively and with greater precision. Only by discarding previous beliefs and replacing them with a different set is the newer model accepted.²

There are mathematical models designed to represent the elements within the structure of an organization and to describe their relationships with the organization’s goals. These mathematical models use equations representing the presumed “reality” to solve “what if” questions by changing the model parameters.³ In this case, the organization under consideration is a school.

Why Build a Model?

According to Williams, the value in model building is threefold.⁴ First, building a model often reveals structures, elements, and relationships usually taken for granted until the underlying assumptions are stated

and tested. Once the original ideas are stated and tested, they usually give way to more sophisticated and accurate representations of the actual situation. Second, once the model is constructed, analyzing it mathematically suggests courses of actions not readily apparent. In essence, the model challenges conventional thinking. Third, experimentation is possible within a model that is not practical in actual situations. Through experimentation more potentially successful options may be identified. Unlike “seat-of-the pants” decisions, models can be tested.

Fundamental Assumptions

To start, there are five fundamental assumptions regarding desirable school outcomes: (1) Student outcomes as measured by achievement tests are appropriate measures of school performance; (2) Other student outcomes, such as school retention, graduation, and employment rates are also appropriate measures of school performance; (3) Because many of the measures of student performance are highly associated with the school’s community socioeconomic status (SES), it must be taken into consideration; (4) Because all schools will not have the same success in achieving student outcomes due to differences in organizational effectiveness, school effectiveness should also be taken into consideration; and (5) When considering alternative policies to achieve the desired outcomes, cost-effectiveness is a critical component.

Next are five fundamental assumptions regarding modeling school organizations: (1) Based on the properties of the normal curve, achievement tests are stochastic in nature, and the model must be consistent with these stochastic properties;⁵ (2) Because achievement tests have a definite upper limit rarely, if ever, achieved by all students within a school, “perfection” is not obtainable, and therefore there is a point after which additional resources will produce diminishing returns; (3) Schools pursue multiple outcomes simultaneously; (4) Schools are complex organizations balancing multiple elements and processes to achieve their multiple goals; and (5) Because there will be a unique solution for each modeled school based on the initial conditions of the organization, there will not be a single policy to achieve the desired results applicable to all schools.

Conditions to Achieve Optimization

Mathematical programming (sometimes called “linear programming”) is merely a method of solving simultaneous equations. The solution could represent the optimal use of resources to produce the optimal level of outcome. The basic structure of a mathematical programming problem is illustrated by this example:

$$\begin{array}{ll}\text{Maximize:} & 3X + 2Y \\ \text{Subject to:} & X + Y < 4 \\ & 2X + Y < 5 \\ & -X + 4Y > 2 \\ \text{Constraints:} & X \geq 0 \\ & Y \geq 0\end{array}$$

Establishing equations accurately representing the organization to be modeled is the key to mathematical programming. These equations must meet certain conditions in order to be solved. The four basic conditions listed below are developed throughout the paper:

(1) There must be a single expression, the “objective function” to be maximized, minimized, or set to a specific value representing the underlying purpose of the model.

(2) There must be simultaneous equations accurately representing the structure and elements of the organization and their relationships

James L. Phelps is a former Education Assistant to the Governor of Michigan and Deputy Superintendent in the Michigan Department of Education.

to the organization's desired outcomes for which there are solutions or boundaries.

(3) The boundaries may be of various types:

- Intersection of lines (lines with positive and negative slopes)
- Maximum or minimum points of nonlinear functions (curves with a change in the sign of the slope)
- Diminishing returns (curves with a changing slope approaching asymptotic).

(4) There are usually constraints or a series of expressions setting limits on any or all of the variables. Cost is a frequent constraint.

Why Is Education Different?

Much of the mathematical modeling has been developed in areas such as business where the outcomes are in discrete and limitless increments, and the relationships are frequently linear. For example, if the purpose of the organization is to produce and sell widgets, it is straightforward to calculate how many machines and how much material is needed and what staffing levels are required to operate and maintain the equipment. The associated cost with these elements can also be determined. With this information, different combinations can be explored to determine the best—the most economical—way to proceed. There is no limit as to the number of widgets that can be produced although there may be a limit to the number that can be sold.

In contrast, there are areas, such as education, where outcomes are stochastic—measured by normally distributed achievement tests—and the relationships among organizational variables and outcomes are less straightforward. The results from a change in the organization's activity can only be estimated based on probabilities and within a margin of error rather than with great certainty. Also, there are definite limits. If the average score on a standardized achievement test were 100, there is no way to modify the school organization at any cost to double the score, to 200, if a perfect score was 150. Indeed, while it is possible to make a plethora of widgets virtually identical, it is virtually impossible to make the achievement of a plethora of students identical.

Given the difference between nonstochastic manufacturing products and stochastic education outcomes, the model presented here is designed to address the fundamental question raised previously: How can schools allocate available resources to best achieve student performance goals?

The Production Function and Regression Analysis

Conceptual Elements of Production

A helpful model for thinking about organizations is the production function.⁶ Conceptually, the production function is divided into three main parts: (1) the outcome to be achieved; (2) the input required; and (3) the process used to convert the input into the outcome. It is represented by the following equation:

$$\text{Outcome} = \text{Input} + \text{Process}$$

In most cases, each of the parts is comprised of many variables.

As the equation requires, the level of outcome increases if either the input or process variables increase, but the “trick” is to determine which input or process variables to increase and by how much. In modeling, if the levels of inputs and process variables and their relationships to the outcome are known, the level of outcome can be predicted. This knowledge provides insights on how a change the input and process

levels will alter the level of the outcome. When deciding the variables to include and the mathematics to estimate the relationships and to calculate the predicted outcome, the basic operational assumptions of the organization, either implicit or explicit, are incorporated into the model.

The production function may be optimized via mathematical programming when the input and process variables and their relationships to the outcomes are known. When the relationships are unknown, they are usually estimated through the statistical model of regression. However, regression analysis does not directly provide answers to optimization questions.

Estimating Relationships Via Regression

The basic regression model estimating the relationships (weightings) is straightforward:

$$\text{Outcome} = X_1 * I_1 + X_2 * I_2 + \dots X_n * I_n + Y_1 * P_1 + Y_2 * P_2 + \dots Y_n * P_n + \text{Unknown} + \text{Error}$$

The X's and Y's represent the estimated weightings measuring the relationship between the outcome and the input and process variables. The I's represent the variables defining the inputs. P's represent the variables defining the processes. “Unknown” represents the important variables in the production function for which data are unavailable. “Error” represents the portion of the equation that cannot be explained because of measurement error.

In order to get meaningful results, the distributions of the outcome, input, and process should be normal or near normal with a substantial degree of variation. Variation is required to accurately place each observation. In education, student achievements test are designed based on these characteristics and, therefore, are stochastic. (See footnote 5.)

Interpreting Regression Results

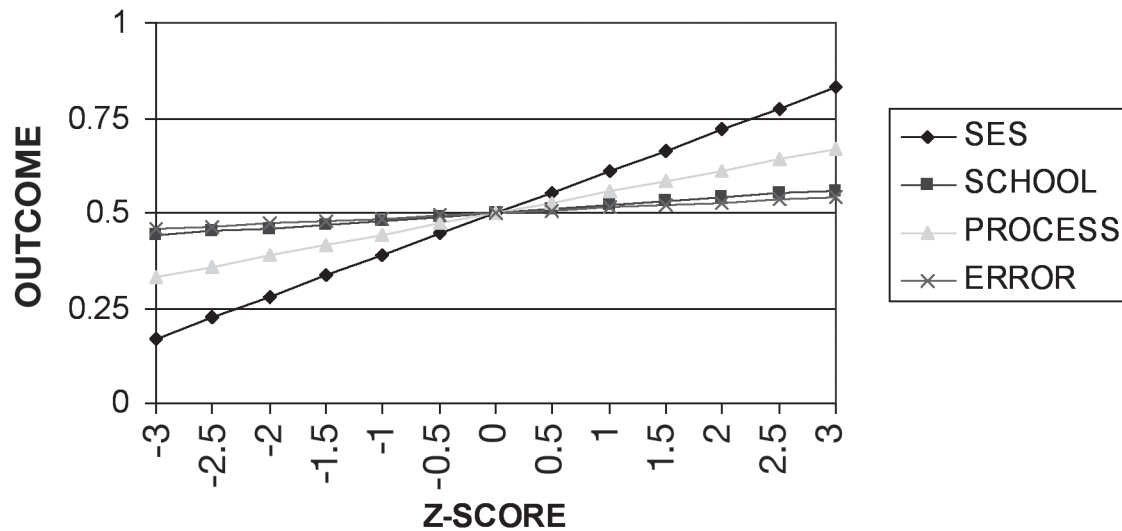
The most common conclusion of a regression analysis is the statistical significance of the weighting; if it is significant, then it is thought appropriate to increase the level of the input or process variable. However, the level of significance does not help determine how much to increase the variable.

The weighting measuring the relationship between the outcome and the independent variable(s) is interpreted as slope; that is, the unit-change in the level of the outcome for each unit-change in the input or process variable. Slope is also the mechanism for predicting the most likely value of an outcome from the known value of an input or process variable. The slope does provide some greater help in determining which variable to increase because it only makes sense to increase the variable(s) with the highest slope—“the biggest bang.”

Many of the following illustrations have been taken from a previous study by the author where the production function was divided into the community input of socioeconomic status (SES) and the school inputs of staffing quantity, staffing quality, and other financial resources.⁷ There were no direct data representing the process, which is usually the case. The process component was defined as the effectiveness of the school organization to produce scores higher than what was predicted from knowing the other inputs—the residual. The slopes of the categories of the study are depicted in Figure 1.

Because each of the variables has a unique descriptive statistic, it is difficult to compare their influence on achievement without first converting all outcomes and variables to standard scores (Z-scores). The slope is then the standard regression coefficient. This is the

Figure 1
Regression Slope



convention in the remainder of this paper. Most frequently the graphic representations of the outcome and variables is based on “Cartesian” geometry with the navigation point being the origin (X and $Y = 0$) with the outcome(s) on the Y -axis and the variable(s) on the X -axis. Because the mean value of an independent variable predicts the mean value of the outcome (dependent variable), charting mean against mean as the navigation point will be used. (A standard score or Z -score of zero is the mean.) The outcome in this illustration is measured in percentiles for reasons to be given later.

With this interpretation of slope, there comes a predicament: Why increase any but the variable with the highest slope if the other variables will make less of a difference in increasing the level of the outcome? This contradicts one of the basic assumptions of the production function: It takes a combination of variables combined in a balanced way to improve outcomes rather than just one or two variables in high concentration. This predicament will be addressed later.

There is another aspect to the regression analysis—predicting the outcome level based on the values of the input and process variables. By substituting the actual values back into the regression equation with the estimated weightings, a predicted level results. The difference between the actual outcome level and the predicted outcome level is the residual, or, an unfortunate name, “error.”

Residual as Effectiveness

The notion of the residual being all error is misleading. An important variable may not have been included in the original equation, and, if it were, the error term would be reduced. Therefore, part of the error term is usually due to a misspecification of the equation, but what if the residuals were compared over several periods of time and there was a tendency for the residuals of each observation to have the same sign and magnitude? In this case, it would be fair to assume the pattern of the residual actually measures something real but unobserved. Because organizations utilize their resources to different degrees of effectiveness, a logical conclusion would be for any consistent pattern of the residuals over time to be associated with an unobserved effectiveness factor.⁸

Limitations of Regression to Optimize

While of great value in estimating the magnitude of relationships, the statistical model of regression does not directly address the fundamental question of how to best allocate resources among the input and process variables.

The basic assumption of the regression model is that of linearity of the weightings; as each unit of the independent variable is increased, there will be a constant increase in the outcome. To have a “perfect” outcome, e.g., all students with a perfect score, it is mathematically possible by increasing any one of the model inputs sufficiently to obtain a predicted perfect score. In practice, this situation does not occur. Indeed, some students achieve perfect scores within the existing resources, but there is a distribution of scores for all the students with the average score well below perfect. In order to achieve a perfect score for an individual school, the variation among students would have to be reduced to zero as well as an improvement of all scores below perfect. Perhaps this could be achieved by eliminating some students from the population or “dumbing-down” the test, but these efforts would negate the basic purpose of assessing student progress. At the heart of the stochastic assumption is the recognition of the existence of individual differences over which the school has only partial control.

While it is possible to introduce some degree of nonlinearity into variables, e.g., introducing an additional term calculated by squaring the variable value, these results are seldom significant. Even if significant, there is seldom a change in the sign of the slope—a maxima or minima point—and thus, predicted “perfection” is still possible.⁹

Thus, if all the variables are linear (or at least always with a positive slope), what is the optimum allocation of resources? Initially this question may be addressed by standardizing the weightings, converting all variables to standard scores so they are comparable. After the weightings are standardized there is the question of cost. This can be addressed by comparing the standardized weightings per dollar.

After these procedures are completed, there is still no answer to the fundamental question. Because only one variable will have the best cost per unit improvement of the outcome defined as cost-effectiveness,

mathematical logic still leads to placing all the resources in a single variable. While logical mathematically, it is not logical operationally. Organizations operate effectively because of blending many variables to achieve the best outcome, not by selecting just one “basket for all the eggs.” In addition, most organizations have the mission of achieving multiple outcomes, but regression, with just a single equation, addresses only one. While various outcomes could be combined to form a single outcome, much of the valuable information unique to each outcome would be lost.

In summary, the regression statistical model as an optimization tool is deficient in four respects: (1) It does not directly model the relationships among multiple outcomes and the organizational inputs and processes; (2) It assumes linearity in the weightings, precluding a systematic balancing of the various variables to achieve the best possible outcomes; (3) With linearity, outcome “perfection” can be achieved given sufficient resources and investment in only one variable; and (4) There is no provision within the model for addressing cost-effectiveness.

Using Regression to Seed an Optimization Model

Based on everyday experience, the assumptions represented by the statistical model of regression are not consistent with school organizational reality. One would be hard-pressed to identify a school organization operating under the assumptions of the regression model, but is it possible to take the analytical results from regression and insert them into a mathematical programming model more consistent with reality?

Estimates from Regression Into Mathematical Programming

Regression, with a single outcome, is not designed to optimize. This can be easily addressed by formulating individual equations for each of the outcomes, establishing a set of simultaneous equations, an essential characteristic of mathematical programming. The explicit goal is to achieve the highest possible level for the sum of the multiple outcomes. (A mathematical transformation can be made to accommodate something like a dropout rate where it is desirable to have the rate low.) If some outcomes were thought to be more important than others, a weighting system among the outcomes could be included. Addressing the second and third deficiencies mentioned above is more involved.

Transforming Relationships to Achieve Diminishing Returns

Conceptually, there are three general ways to describe the relationship between inputs and outcomes, sometimes called “returns to scale”: (1) Increasing returns to scale or the inverse, decreasing returns to scale; (2) Maxima or the inverse minima; and (3) Constant returns to scale. (See Figure 2.) Note that one curve is increasing for the first half and decreasing the second. The slope determines the type of relationship based on whether the slope is increasing or decreasing, whether there is a point where the slope is zero, or whether the slope is constant. The return is measured in percentiles.

In order to solve simultaneous equations, as mentioned previously, there must be either intersection of lines; maxima or minima points of curves; or curves representing diminishing returns. Assuming positive linearity of each regression weighting, there can be no intersection of lines or maxima and minima points, therefore no solution to the equations. The most likely alternative to solving the simultaneous equations is to form nonlinear functions indicative of diminishing returns.

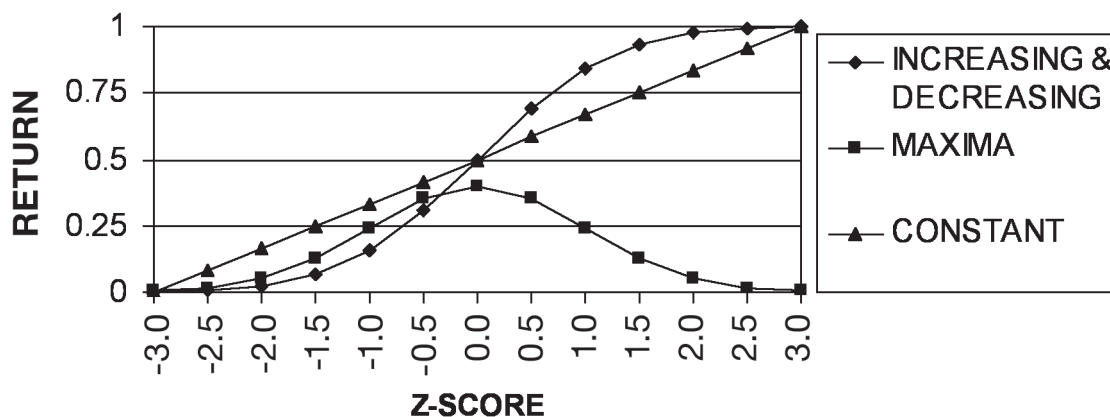
Diminishing Returns Function Within Regression Analysis

At this point, there is an essential digression to demonstrate mathematically the existence of a nonlinear function indicative of diminishing returns based on regression analysis.

Students in beginning statistics courses are taught several descriptive statistics, but they most likely do not fully appreciate their full beauty and power. Usually, an early step is to construct a histogram underlying the distribution of a bell-shaped curve. Students are then asked to calculate the mean and standard deviation. After calculating the mean, the deviations from the mean are calculated, these deviations are squared, and then they are summed. The result is called the sum of the squares and commonly noted as “SS.” The sum of the squares is then divided by the number of observations (N) to produce the mean of the squares (MS). This is also called the variance as symbolized by σ^2 . When the square root of the variance is taken, the result is called the standard deviation or σ . The variance is some notion of area, but area of what? The standard deviation is some notion of length, but length of what?

The primary purpose of regression analysis is to make predictions regarding the level of the dependent variable (outcome) based on the values of the independent variables (inputs). The basic idea is to plot

Figure 2
Returns to Scale



the dependent variable on the Y-axis and the independent variable on the X-axis to determine if these points tend to fall on a line. While this can be inspected visually, it can be measured with great accuracy mathematically. The line is considered the “best fit” when the distance from the observation point to the regression line is squared, summed for all observations, and minimized. This method is called the “least-squared” solution. The line is represented algebraically as the slope of a line. It is presented in two forms, one using the original values, i.e., the regression coefficient, and another using standard scores—the standard regression coefficient. When the variables are measured in standard scores (Z-scores) and the slope is measured in terms of the standard regression coefficient (r), the value of the outcome can be predicted from the value of the independent variable with the equation: $Z_{(y)} = r Z_{(x)}$.

However, the regression analysis provides another estimate, the amount of variance explained by each of the variables. Regression programs calculate the sum of the squared deviations for the independent variable(s) and well as for the residual, what is not accounted for by the independent variable(s). These sums of the squared deviations are converted to percentages of the total and called the coefficient of multiple determination, or R^2 . It is a measure of the “goodness” or “strength” of the prediction of the variable(s), with the higher value indicating a greater strength. When the R^2 is 100%, there is “total strength,” and when the R^2 is 0%, there is “no strength.” When the percentage of what can be explained or attributed is added to the percentage of what cannot be explained or attributed, the sum is 100%.¹⁰ Can the R^2 be related to the probability curve?

What Is the Probability Curve?

The idea of the probability curve is rather straightforward. If one tossed a number of coins a number of times and calculated the number of times each combination of heads and tails occurred, the result would form a histogram high in the middle and low at the edges. (The probability of each combination can be calculated via a binomial expansion and represented by the coefficients depicted in Pascal's triangle.) The probability curve is merely the probability histogram as the number of observations approaches infinity and converted to a continuous bell-shaped curve. It answers questions regarding the

probability that any event will occur. Of course, there are limits or boundaries to probability. No event can occur more than 100% or less than 0% of the time.

The continuous bell-shaped curve is represented by the expression, $e^{-Z^2/2\sigma^2}$. The denominator of the exponent contains the variance (σ^2) from the descriptive statistics. The area under the probability curve, when normalized, is by definition 1 because the chances of something happening cannot be greater than 100 percent; so there must also be a denominator added to the expression representing the area of the curve. When the denominator equals the area of the numerator, the result is 1. The area of the probability curve is $\sigma\sqrt{2\pi}$, so the complete expression for the normalized curve is $(1/\sigma\sqrt{2\pi}) (e^{-Z^2/2\sigma^2})$. The standard deviation (σ) appears in the calculation of the area. The variance and standard deviation are parameters of the probability curve.

Reformulating the Regression Results Into the Normal Curve

From regression, the explained variance by the independent variable plus the unexplained variance equals 1, as represented by the following equation:

$$R^2 + K^2 = 1$$

R^2 is the explained variance, and K^2 is the unexplained variance. If additional variables are added to the equation, the proportional relationship is maintained as represented in the equation:

$$R_1^2 + R_2^2 + K^2 = 1$$

Therefore, each term in the equation explains a proportion or percentage of the total variance. Variance is a measure of area based on the principle of squared deviations.

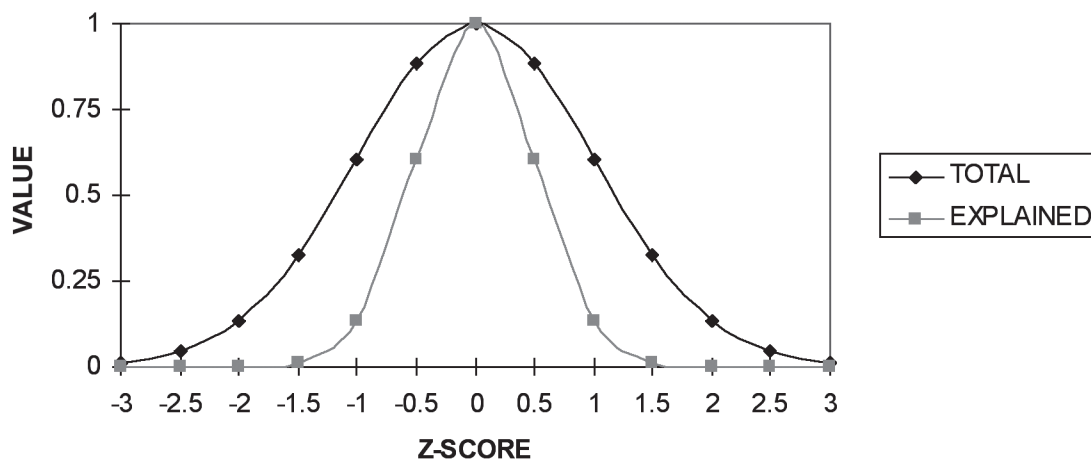
For the ease of notation, I will call the area of the probability function $f(z)$, where the measurement of the X-axis is in terms of Z-scores, or standard scores, and the area of the probability curve is normalized ($f(z) = 1$), and is represented by the following equations:

$$(R_1^2 + R_2^2 + K^2) f(z) = 1$$

or

$$R_1^2 f(z) + R_2^2 f(z) + K^2 f(z) = 1$$

Figure 3
Comparison of Variance



Given a specific observation as measured by a Z-score, the relative position of that observation can be easily calculated and reported as the percentile ranking. Therefore, the predicted placement, measured as a percentile ranking $Y(p)$, for a specific observation across all terms is calculated by substituting the appropriate Z-score for each term, with K^2 representing the margin of error, as follows:

$$Y(p) = R^2_1 f(z_1) + R^2_2 f(z_2) + K^2 f(z_3)$$

or

$$Y(p) = R^2_1 f(z_1) + R^2_2 f(z_2) + \pm 1/2 K^2 f(z_3) \text{ Q.E.D.}$$

In other words, the reformulated equation is a regression equation measured in terms of the proportion of area under the normalized curve or percentile and the predicted outcome value can be calculated for any combination of Z-scores. This representation of the R^2 is easily demonstrated graphically for it now relates to the proportion of area under the normal curve. (See Figure 3.)

Interpretation of the Normal Curve

While there is a maximum point at a Z-score of zero (the mean), the slope then turns negative, signifying declining returns rather than the more plausible diminishing returns. There is no evidence or theory suggesting that benefits would or should start decreasing when resources move past the mean. Is there another way in which to view these curves that is more consistent with evidence and theory?

To review, the area under probability curve ($\sigma\sqrt{2\pi}$) is determined by the width parameter σ (standard deviation). The probability curve is represented by the expression $e^{-Z^2/2\sigma^2}$. The Z symbol Z represents the standard score or Z-score, and when Z equals zero, the function equals one. (See Figure 3.) As one might expect, the calculations of area of this expression are messy, to say the least. Instead, a single ideal normalized curve is established: area = 1 when $\sigma = 1/\sqrt{2\pi}$. The calculations of area are made on the ideal curve and given either in a table in a statistics book or as a part of a computer program. Hence, the cumulative area under the normal curve can be calculated for any given Z-score.¹¹ The formal name of the resulting S-shaped curve is the standard normal cumulative distribution, or cumulative area curve for short. Given this metric, it is possible to determine easily the percent of observations above and below a given score—the percentile.

This cumulative area curve represents the concept of diminishing returns because the benefits gradually reduce as the variable increases but never reaches a maximum point. (See Figure 4, marked "Area.") This representation appears to match the evidence and theory of the correlates of student performance. One could argue that having more

textbooks in the classroom would be positively related to student outcomes, but only up to a certain point. After each student has one textbook, what would be gained by having more? Even in the case of class-size, it would seem illogical to argue that more than one teacher per student at any one given time would lead to higher achievement than having just one. A case can be made in virtually all circumstances that there is a point where additional resources would reap little or no benefit. Optimization will help determine where these points lie.

Importantly, the cumulative area curve can be used for solving simultaneous equations. Even more importantly, the shape of the cumulative area curve is determined by the R^2 value from regression analysis. The probability and cumulative area curves are related through the mathematics of calculus. The cumulative area curve is the integral of the probability curve and the probability curve is the derivative of the cumulative area curve. This means the probability curve is the slope of the cumulative area curve at the same Z-score. At a Z-score of zero, the value of the probability curve is one, so the slope of the cumulative area curve is also one. When area curve is adjusted for the R^2 value, the slope of the curve at a Z-score of zero is the R^2 value. Through the application of mathematics, the estimates from regression analysis can be transformed into a function suitable for solving simultaneous equations.

By way of illustration, if there were a single independent variable in the equation and the R^2 was 1.00, there would be a perfect relationship between the independent variable and the outcome. The key is that the distribution of the independent variable is measured in terms of standard scores, or Z-scores while the outcome or dependent variable is measured in terms of the proportion of variance explained—the cumulative area under a probability curve, or percentiles. For every standardized-unit increase (Z-score) in the independent variable, there is a corresponding increase in the outcome. In graphic terms, the distribution of the independent variable moving from the lowest to the highest corresponds with the cumulative area under the curve of the outcome from lowest to highest. In other words, the distributions of the outcome and independent variable would be identical but measured in different terms, and, thus, the independent variable explains all the variance of the dependent variable. (See Figure 5.)

If the R^2 were zero (.00), there would be no relationship between the independent variable and the outcome. There would be no width to the outcome variable distribution and no width to the cumulative outcome distribution. In essence, every value of the independent variable would make the same predicted value for the outcome—the mean value. Instead of a spread of the cumulative distribution, there would be a single horizontal line at the mean (50th percentile). Thus,

Figure 4
Normal Curve

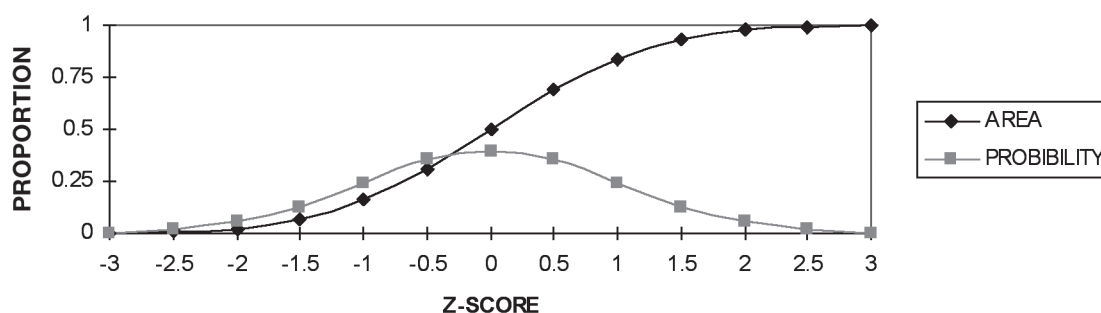
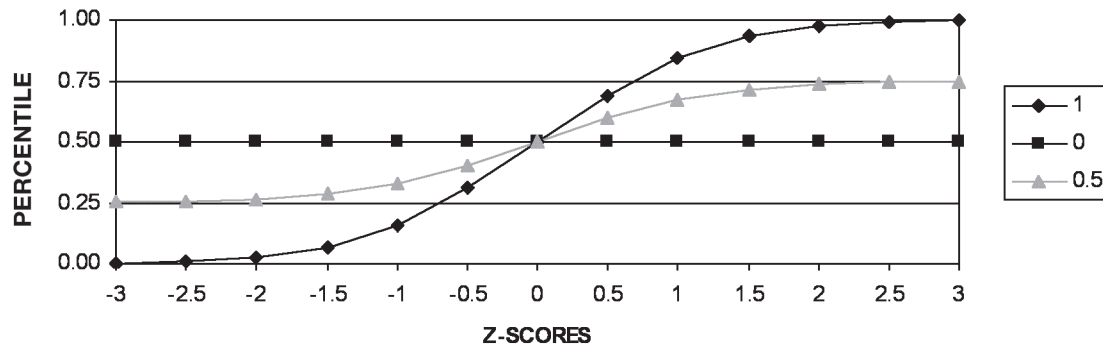


Figure 5
Area Under Curve



the independent variable would explain none of variation in the dependent variable, and the slope of the area curve would be zero. (See Figure 5.)

If the R^2 were .50, there would be a strong relationship between the independent variable and the outcome. The mean value of the variable would still predict the mean value of the outcome, but what about the other values? Because the area of the independent variable would be half of the outcome, half of that area (or one-quarter) would be above the mean and half would be below. When graphed, the S-shaped cumulative curve will be asymptotic to lines representing .75 and .25 of the area. These parameters conveniently represent percentiles. (See Figure 5.)

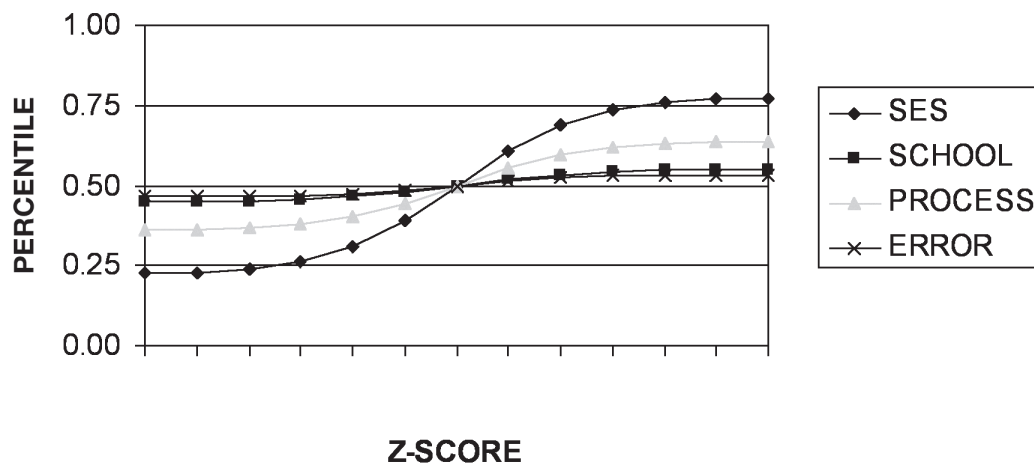
The R^2 terms can be calculated using the respective regression coefficient (r) and the standard regression coefficient (βr). In one sense, this calculation is more precise because it can be negative if r is negative, indicating an inverse relationship between the outcome and the independent variable. On the other hand, a negative R^2 term will not satisfy the summation to 1.0 and is changed to a positive (absolute value) for that purpose in statistical programs. This anomaly should be considered when determining the value of R^2 in a model. A negative coefficient makes the same contribution to the explanation of an outcome as does a positive value, so if there is an inverse relationship between the independent and dependent variables, the sign of the R^2 value should be set to negative in the simultaneous equations.¹²

In summary, the relationship between the distribution of a probability curve and the cumulative area curve is a straightforward transformation suitable for solving simultaneous equations. Conceptually, it is merely converting the outcomes to percentiles and the independent variables to standard scores.

The S-shaped curves are all asymptotic to the lowest and highest values as determined by the R^2 , thus solving the boundary dilemma of achieving perfect scores by allocating an infinite amount of resources. While an increase of resources may improve the outcome level, it is both conceptually and mathematically impossible in this interpretation to achieve perfection because the asymptotic curve will never reach the maximum. This situation is consistent with the basic assumption of school performance. When applied to actual estimates of the production function, the respective relationships are depicted in Figure 6.¹³

With this transformation, the mechanics of optimization are rather straightforward even though the preparation of the data is somewhat tedious. The multiple R^2 weightings are inserted into a set of simultaneous equations based on the cumulative area function. Then, the principles of mathematical programming are applied to solve for the optimal levels of variables that will produce the highest level of summed outcomes. Importantly, the simultaneous equations model also requires the inclusion of constraints consistent with organization practice, the most notable being cost. Other upper and lower limits can

Figure 6
Relationships



be included as organizational practice requires. It should be emphasized that this solution is not for the weightings as they were estimated via regression analysis in the form of the R^2 . Rather, the solution is for the values of the independent variables that will predict the best result—the highest predicted level of outcomes summed across the several equations.

The shift from the standard regression model to an optimization model may be more difficult psychologically than mathematically. Because of common use, most people are more comfortable with regression, but the critical difference is in the acceptance of the deficiencies listed above and their practical consequences. It is much easier to believe in continuous improvement for increased resources than it is to believe in diminishing returns—a point where an increase in resources produces little, if any, improvement. However, can the simultaneous equations with the transformations actually be solved and will the solution provide insights into the fundamental question—what is the best allocation of resources to achieve the optimal outcomes?

The Optimization Model

The optimization model takes a form common in mathematical programming, with the following elements: Objective function as the sum of the outcomes; equations defining the relationships between multiple variables and the outcomes; equations calculating the cost; and constraints limiting the upper and lower bounds of the variables

There is no method to predict future outcomes with complete accuracy. There are changes in the organization plus there is a certain degree of measurement uncertainty. As a result, the estimated outcomes are stochastic and based on predictions. Therefore, there must be two sets of simultaneous equations defining the outcomes, somewhat like a “before” and “after.” Before and after are not different time periods; rather, they are the predicted results before and after the optimization. Before estimates the actual predicted target utilizing the existing variable values, and after estimates the optimized predicted target utilizing the optimized values.

The basic structure of the equations is similar in form to regression equations:

$$\text{Outcome}_a = W_{a1} * V_1 + W_{a2} * V_2 + \dots W_{an} * V_n + \text{Residual}$$

$$\text{Outcome}_b = W_{b1} * V_1 + W_{b2} * V_2 + \dots W_{bn} * V_n + \text{Residual}$$

$$\text{Outcome}_n = W_{n1} * V_1 + W_{n2} * V_2 + \dots W_{nn} * V_n + \text{Residual}$$

W 's are weightings, potentially different in each equation while V 's are variables, the same in each equation. For each set of equations, the outcomes are summed to produce a target:

$$\text{Actual Predicted Target (Before)} = \text{Set One (Outcome}_a + \text{Outcome}_b + \dots \text{Outcome}_n)$$

$$\text{Optimized Predicted Target (After)} = \text{Set Two (Outcome}_a + \text{Outcome}_b + \dots \text{Outcome}_n)$$

The objective function, the value to be maximized, is the gain in the predicted outcomes achieved by changing the resource allocation pattern:

$$\text{Objective Function} = \text{Optimized Predicted Target (After)} - \text{Actual Predicted Target (Before)}$$

The constraints control the total cost as well as minima and maxima for each of the variables:

$$\text{Total Cost} = V_1 * \$1 + V_2 * \$2 + \dots V_n * \$n$$

$$\text{Total Cost} = \text{specified value}$$

$$V_n \geq \text{specified value}$$

$$V_n \leq \text{specified value}$$

The weightings (W 's) are the R^2 for the respective variables (V 's). The R^2 are estimated via regression analysis, but, as noted earlier, it can be negative. The respective variables (V 's) are the actual observation values for calculating the actual predicted target in the first equation, and the optimized values for calculating the optimized predicted target in the second equation. The total cost of each set of equations is calculated by multiplying the value of the variable (V) by the average cost of the variable (\$).

There must be a cost constraint; organizational resources are always limited. The values of $V_1 * \$$, $V_2 * \$$... $V_n * \$$ when added must be equal to or less than a specified amount, the total cost of the resources available to the organization. The purpose of optimization is to maximize the sum of optimized predicted outcomes while staying within the cost boundary. A “cost” expression is inherent to optimization but missing in regression.

What makes this model unique is the function used to represent the relationship between the independent variables and the outcomes. Because achievement outcomes are measured in stochastic terms—normal distributions—the relationships are measured in the same way. Rather than defining the relationship between the independent and dependent variables as the linear slope, the relationship is defined in terms of a type of standard scores. Because the area under the normal curve can be represented in terms of percentiles, the unique function is the integral of the normal curve—an S-shaped curve—adjusted by the degree of relationship, the R^2 . The higher the R^2 , the more vertically expanded the S-shaped curve, and vice versa. The integral of the normal curve is asymptotic at high and low points, so it is impossible to reach the absolute maximum or minimum points. While the slope at the mid-point (Z-score of zero or the 50th percentile) is the R^2 value, the slope gradually diminishes as it progresses upward and is symmetrical downward. (See Figure 4.) The basic idea is to increase the allocation level in favor of the variable when the slope is the greatest and decrease the allocation level in disfavor of the variable when the slope is the least. This decisionmaking rule is the essence of diminishing returns.

Data Requirements

Most state departments of education have data on the most frequently considered variables, such as the numbers of staff, salaries, qualifications, etc. The model can be specified for either school districts or school buildings. There is the obvious relationship between the sophistication of the data and the model; that is, the more sophisticated the data, the more sophisticated the model will be. With advancements in technology, the data for the model are easily obtainable through information systems.

The following data are required for the model: (1) Population data on the outcomes and variables to calculate means and standard deviations; (2) observation data for the outcome and variables, including actual levels; (3) cost data for the variables of the observation; and (4) estimates of the relationships between the outcomes and the various variables in terms of the R^2 .

The model can be established based on two types of scenarios: Improvement based on redistribution of existing resources when the constraint of total cost is set at the existing level (an increment of

Table 1
Summary Table of Data for Elmstown School and State

		Staffing					
		Students	Classroom Teachers	Support Staff	Teacher Aides	Administrators	Sum
State	n (total)	100,000	4,000	1,000	750	650	6,400
	Per thousand		40.00	10.00	7.50	6.50	64.00
	Std Devition		5.00	2.00	2.00	2.00	11.00
School	n (total)	1,000	40	10	8	7	64
	Per thousand		40.00	10.00	7.50	6.50	64.00
	Z-Score		0.00	0.00	0.00	0.00	
	Percentile		0.50	0.50	0.50	0.50	
		Salaries					
			Classroom Teachers	Support Staff	Teacher Aides	Administrators	Sum
School	Mean		\$50,000	\$55,000	\$25,000	\$75,000	
	Total		\$2,000,000	\$550,000	\$187,000	\$487,500	\$3,225,000
		Student Achievement and Socioeconomic Status					
		Math3	Math5	Read3	Read5	SES	
State	Mean	1,400	1,200	1,400	1,300	50	
	Std Deviation	300	200	350	250	10	
School	Z-Score	0.00	0.00	0.00	0.00	0.00	
	Percentile	0.50	0.50	0.50	0.50	0.50	
		Effectiveness					
		Math3	Math5	Read3	Read5		
State	Mean	0.00	0.00	0.00	0.00		
	Std Deviation	0.06	0.06	0.06	0.06		
School	Mean	0.00	0.00	0.00	0.00		
	Z-Score	0.00	0.00	0.00	0.00		
	Percentile	0.50	0.50	0.50	0.50		

zero); or improvement based on a cost increment when the constraint of total cost plus an increment is set.

An Optimization Example

The optimization model is illustrated here using fictitious data from a state and a school building—Elmstown. The purpose of the optimization is to improve the predicted achievement outcome levels by changing the staffing levels in the categories of classroom teachers, support staff, teacher aides, and administrators. For the state data, converting each variable into “staff per one thousand students” normalizes the raw numbers. The means and standard deviations are required in order to calculate Z-scores and percentiles. Also, the mean and standard deviation are required for each of the outcome variables, in this case mathematics and reading at the third and fifth grades, in order to

calculate Z-scores and percentiles. The same statistics are required for SES and effectiveness variables for each of the outcomes.

At the school building level, data are required for the number of staff in each category as well as the average salary for each staffing category. With this data, the salary total is calculated (number of staff times the average salary summed across categories). Using the state data, Z-scores and percentiles are calculated for the achievement variables. These data are seeded into an Excel spreadsheet to carry out the optimization. In order to focus on the school input variables, SES and effectiveness variables are set to the mean, or 50th percentile. In an actual example, these data will assure the analysis optimizes the school variables without the influence of the other factors. Table 1 illustrates the state and school data.

Setting Model Parameters

In order to carry out the optimization, two sets of parameters must be added. These estimates do not have to be exact, but do have to fall within a reasonable range. According to Schrage, “The first law of modeling is don’t waste time accurately estimating a parameter if a modest error in the parameter has little effect on the recommended decision.”¹⁴ The first set of parameters includes the estimates of the relationships between the staffing categories and the multiple outcome variables as measured in terms of the R^2 , the proportion of variance explained by each of the staffing variables. The researcher selects these estimates based on ranges produced by regression analysis of the population. There is, however, a mathematical limit to these estimates: The sum may not exceed 1.00. The second set of parameters contains the minimum and maximum levels for each of the staffing categories. These constraints address other practical considerations required by the organization and are selected by the researcher. There also must be a cost constraint, the total amount available to spend.

Calculations in the Equations

The model contains two sets of equations predicting the outcomes before and after the optimization. The before scenario is based on the actual organization values—the predicted target—and the after is based on the optimized values—the optimized target. The calculation for each of the terms (variable times weighting) is particularly noteworthy. The calculation is based on the notion that the best predictor of an outcome is the mean (Z-score = 0, or 50th percentile) when no other information is available. So when some information is available, the calculation is measured by how much the estimate varies above or below the 50th percentile. The calculation for each term is as follows:

$$\text{Term} = R^2 * (\text{Percentile} - .5)$$

The predicted outcome is the sum of the terms plus the 50th percentile. The calculation answers the question: How many percentiles above or below the 50th percentile will the prediction be? The calculation is as follows:

$$\text{Outcome} = \sum \text{Terms} + .5$$

The optimization process selects new values for each of the staffing categories producing the optimal gain above the predicted target, also known as (the objective function or “gain in target,” given several constraints. In this illustration, the major constraint is the total cost of staffing, which must be the same for the before equations and the after, or optimized, equations. Of course, the conditions of maximums and minimums for the respective variables in both equations must be honored. In essence, this scenario is to redistribute the existing financial resources across the staffing categories. If the total cost of the optimized equations were set higher than the before cost, the scenario would be incremental in nature. In Excel, the solver identifies the objective function as the “target cell” and optimum values as “by changing cells.” The constraints are identified in under the heading, “subject to the constraints.”

Because the optimization is conducted here on a single observation—here a school building—the solution is unique to this building. The regression model implies the same outcome increase for the same change in variable level for every observation regardless of starting point. In contrast, the optimization depends on the unique starting points of each observation, so the amount of increase is always unique.

In order to make the results of the two predicted outcome values as close as possible to the actual outcome values, it is critical to include SES as a variable in the model. In virtually all studies, SES is the highest predictor of student achievement.¹⁵ A measure of school effectiveness is also included to make the predictions as robust as possible.

Return to the Production Function

Earlier, the notion of the production function was introduced. The original conceptualization was:

$$\text{Outcome} = \text{Input} + \text{Process}$$

At this point in the discussion, it has more practical implications. Through the refinement process, the function has become more sophisticated. First, the input has been divided into two categories, the school inputs and the community input of SES. Second, the process element has taken on the character of the effectiveness variable represented by the regression residual. The residual of a regression equation is comprised of an unobserved variable, a variable not in the equation, and error due to the inaccuracies in measurement. Assuming the residual is an unobserved variable of effectiveness, it can be separated from the error by averaging the residual over time. The average is the effectiveness portion, and the difference between the average and the residual is the error. The production function evolves into the form:

$$\text{Outcome} = \text{SES} + \text{Effectiveness} + \text{School Inputs} + \text{Error}$$

For the sake of illustration, assume the SES and Error terms are identical over two periods of time. The function express in terms of change (Δ) is then:

$$\Delta \text{Outcome} = \Delta \text{School Inputs} + \Delta \text{Effectiveness}$$

Consider the following scenario. What if the school input weightings in the optimization are inflated or raised higher than what might be considered reasonable? The predicted optimized target will then increase, but what if the actual outcome level does not increase at the same pace? The equation demands balancing, so effectiveness declines. Simply stated, within the rigors of the mathematical model, any overstatement of school inputs will be offset by an decrease in the level of school effectiveness. Hence, attempts to “game the system” by inflating inputs will have the consequence of being labeled less effective.

Table 2
Range of Weightings

Subject	SES	Effectiveness	Error
Math3	0.532	0.381	0.087
Math5	0.635	0.297	0.068
Reading3	0.712	0.223	0.065
Reading5	0.706	0.226	0.068
Mean	0.646	0.282	0.072

Table 3
Optimization Results

Table 3.1 Original Values and Optimal Values										
	SES	Effectiveness	Classroom Teachers	Support Staff	Teacher Aides	Administrators	Total Cost	Increment		
Original Values	50.00	0.00	40.00	10.00	7.50	6.50				
Percentile	0.50	0.50	0.50	0.50	0.50	0.50				
Cost			2,000,000	550,000	187,500	487,500	3,225,000	100,000		
Optimized Values	n/a	n/a	44.42	5.00	10.63	7.51	3,325,000*			
Change			4.42	-5.00	3.13	1.01				
Z-Score			0.88	-2.50	1.57	0.51				
Percentile			0.81	0.01	0.94	0.69				
Cost			2,220,846	275,000	265,806	563,348	3,325,000*			
Minimum			35	5	5	3				
Maximum			50	15	15	10				
*Must be equal										
Table 3.2 R-Square with Goal										
	SES	Effectiveness	Classroom Teachers	Support Staff	Teacher Aides	Administrators	All School	Total	Error	Total
Math3	0.600	0.2500	0.030	0.020	0.010	0.020	0.080	0.930	0.070	1.000
Math5	0.600	0.2500	0.035	0.020	0.010	0.020	0.085	0.935	0.065	1.000
Reading3	0.650	0.2000	0.035	0.020	0.020	0.010	0.085	0.935	0.065	1.000
Reading5	0.650	0.2000	0.030	0.020	0.020	0.010	0.080	0.930	0.070	1.000
Average	0.617	0.233	0.033	0.020	0.013	0.017	0.083	0.933	0.067	1.000
Table 3.3 Predicted Target										
	SES	Effectiveness	Contribution				School	Predicted	Actual	Efficiency
Math3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	50.00%	50.00%	0.00%
Math5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	50.00%	50.00%	0.00%
Reading3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	50.00%	50.00%	0.00%
Reading5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	50.00%	50.00%	0.00%
Sum	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000			
Average	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	50.00%	50.00%	0.00%
Table 3.4 Optimized Target										
	SES	Effectiveness	Contribution				School	Predicted		
Math3	0.0000	0.0000	0.0093	-0.0099	0.0044	0.0039	0.0078	50.78%		
Math5	0.0000	0.0000	0.0109	-0.0099	0.0044	0.0039	0.0093	50.93%		
Reading3	0.0000	0.0000	0.0109	-0.0099	0.0088	0.0019	0.0118	51.18%		
Reading5	0.0000	0.0000	0.0093	-0.0099	0.0088	0.0019	0.0102	51.02%		
Sum	0.0000	0.0000	0.0405	-0.0395	0.0265	0.0116	0.0391			
Average	0.0000	0.0000	0.0101	-0.0099	0.0066	0.0029	0.0098	50.98%		
Gain in Target	0.0000		0.0405	-0.0395	0.0265	0.0116	0.0391	0.98%		

Ranges of Relationship Weightings

There is no fixed set of weightings measuring the relationship between the outcome and the model variables. Every study will produce different estimates. Nevertheless, most studies fall within some consistent range. The author has not completed a thorough study to document these ranges, but based on data from one state, these ranges, measured in terms R^2 or percentile points, seem to be justified. (See Table 2.) In this state, the influence of SES tends to be about 10 points higher for reading than for mathematics while the influence of effectiveness tends to be about 10 points higher for mathematics than for reading.¹⁶ Each investigator will have to determine a range based on what data are available for the population under study. The consequence of overestimating has already been addressed.

After the data have been entered into the spreadsheet model and the optimization conducted, the results can be presented in a format illustrated by Table 3.

Summary of Results and Analysis

All the school variables in this illustration were set to the mean to more easily focus on the features of the optimization. Therefore, the predicted target and actual outcome levels were all at the 50th percentile. In a real situation, these variables will reflect the actual status of the school. When the optimization is applied, the optimized values are indeed changed in that there is an increase in the more cost-effective variables and a decrease in the less cost-effective variables. The total cost of the pre-optimization and post-optimization is equal, thus an incremental scenario. There is an incremental value that could be set to zero by the researcher for a redistribution scenario. The constraints have been met in that the support category is at the minimum. The gain in the predicted gain in target is an average of .98 percentiles.

The optimization also produces some analytical information of potential usefulness. The contribution of each of the variables for each of the outcomes is provided indicating the respective cost-effectiveness. The average contribution for each of the variables is also provided. The contributions of the school variables are provided separately. There is a check of the R^2 sum to assure that it is not greater than 1. The sum of the R^2 of the school input terms is provided to assure it falls within a reasonable range. A measure of efficiency is given, calculated as the difference between the optimized predicted outcome and the actual outcome level. It could be considered error or doing better (or worse) than predicted. As this example demonstrates, there is a mathematical solution to the stochastic simultaneous equations model. Only by building and interrogating a “live-data” model with all of the policy relevant variables will it be known if there is a practical policy solution.

Observations Regarding the Optimization Model*Modeling through Estimates*

There will never be enough comprehensive and accurate data. Realistically, data can be used to make estimates of relationships between outcomes and input variables; however, these estimates will always vary over time and populations. Importantly, this optimization model is most effective when realistic ranges of the relationships are examined. Because the cost of a variable is known with great accuracy, it is logical that there is an implied relationship between the cost and the cost-effectiveness of the variables. That is to say, if variable A is three times as costly as variable B, then variable A must be three times as effective for the two variables to be equally cost-effective.

Setting the relationship variables first produces the predicted target level. Importantly, the higher the relationship, the higher the predicted target values. This is not a “freebie,” in that the actual relationship values are, by definition, set so half of the observations will do better than predicted and half will not. This difference is in small part due to error in data measurement, but mostly the difference is due to the inescapable fact that some organizations are more effective in turning resources into outcomes. Therefore, if the relationship variables are set too high, indicating that more resources will produce higher predicted outcomes, it will also tend to increase the gap between the predicted outcome and the actual outcome, indicating a higher degree of ineffectiveness. Increasing the relationship coefficients will have the effect of indicating higher potential achievement scores for greater resources, but it will also render the school less effective when the actual results are measured and the school fails to meet the prediction. In essence, the greatest value is achieved when the parameters are set realistically rather than quixotically.

Inevitable Conclusions

As outlined above, there are some inevitable conclusions associated with the optimization model as compared with the regression model. First, because of the inherent nonlinear structure of the optimization model, it is impossible to achieve entirely the desired goal unless the goal has been completely achieved by other similar organizations. That is to say, it is impossible to set values predicting a perfect outcome score unless it has been actually achieved by other organizations, and the Z-score for that organization can be identified. In terms of student achievement testing, it is highly unlikely any organization records perfect scores for all students.

Second, there is an inherent point of diminishing returns due to the nonlinear stochastic function. At a certain point, any given variable will have reached its potential, and investments in other variables will indicate better results. As a general rule, if an organization is among the highest on a given variable when compared to other organizations, an increase the variable will indicate little increase of outcome in the model. On the other hand, an increase in a variable for which the organization is low as compared to others will indicate a larger increase of outcome. Of course, the variables must be compared based on the cost-adjusted value.

Third, as suggested by point two, the solution to the model will be different for each organization, because the starting point is unique to each organization. Theoretically, if all organizations were moved to the high end (for example, the third standard deviation above the mean) for all variables, the predicted results for all organizations based on the allocation of resources would be similar. Any differences in predictions would be based on variables not included in the resource allocation category such as socioeconomic status or effectiveness. In other words, achievement equity is not possible solely through resource allocation. For complete outcome equity, resources, SES, and effectiveness must all be equal.

The optimization model has two basic strategies: (1) Invest in high cost-benefit variables where the organization level is low compared to other organizations; and (2) Do not invest in low cost-benefit variables where the organization level is high compared to other organizations.

Ranges of Input and Process Categories

In supplying the estimate of weighting in the equations, these conditions must be recognized. First, there is a maximum of an R^2 of 1.00. Second, if the estimated weightings are larger than the actual weightings, the effectiveness ratings of the observations will be reduced; that is, the actual performance on outcome will be less than the predicted outcome level. In theory, the weightings will be close to correct when the effectiveness of all observations is normally distributed with a mean of zero. Over the last several decades, educational research has identified several categories thought to be associated with student learning outcomes. The community and school inputs are: SES; staffing quantity (ratio of various staff classifications to students); (3) staff quality (qualifications, experience, etc); and (4) materials and supplies. Less attention has been paid to the process categories of instruction, including time, curriculum, out-of-school influences; and effectiveness. A comprehensive model could include all these independent variables as long as there are data defining the variables and statistics estimating their relationships with outcomes. While outcomes are usually defined by student achievement measures, other desirable outcomes such as dropout rates and college-bound rates could be included in the model as long as the data for the variables and estimates of the relationships are available. Because there tends to be a high degree of correlations among school variables, adding variables to the model does not always have the effect of increasing the predicted levels of the outcomes. Instead, adding variables merely redistributes the influences. Also, because of the correlation between some school variables and SES, it is appropriate to test the model within reasonable ranges.

Testing the Model

There are some elements of school operations for which there are no estimates of the relationship with outcomes. Probably the best example is that of the school year. Mostly because of state laws, virtually all schools are in operation for the same amount of time. Because there is little variation, there can be no estimated relationship in a regression analysis. But there are options within the optimization model. First, the cost of an extension of the school year can be calculated. Second, the cost can be compared with the cost of other options where the relationship with outcome is estimated. With this information, a calculation can be made as to the relationship level of extending the school year to make an equal contribution as the other option. In a more ideal situation, a national or state research initiative could be conducted by first applying the optimization model and then applying an experiment—in this illustration, a longer school year—to determine if the estimates in the model are realized. Surely this is a more practical method than instituting a statewide policy without any experiment evidence.

Sensitivity Analysis

There is a notion of opportunity cost developed by accountants. Simply, it is how much profit can be gained by increasing production by a given amount. In the optimization illustration, a marginal cost-benefit is provided for each element within the model indicating how much would be gained in student outcome by a certain investment. Obviously, it would be appropriate to invest in the element with the highest cost-benefit. However, the cost-benefit will not be the same for each school because each school has a unique starting point.

Summary, Research, and Policy Issues

The model used for investigating school resource allocation questions has a definite influence on the policy conclusions reached. At the beginning of this article, three potential benefits of building a model were identified. First, building a model often reveals structures, elements, and relationships usually taken for granted until the underlying assumptions are stated and tested. Once the original ideas are stated and tested, they usually give way to more sophisticated and accurate representations of the actual situation. Second, once the model is constructed, analyzing it mathematically suggests different courses of actions not readily apparent. In essence, the model becomes a challenge to conventional thinking. Third, experimentation that is not practical in actual situations is possible within a model. Through experimentation, more potentially successful options may be identified. In essence, models can be tested, unlike “seat-of-the pants” decisions. Now it is time to assess if any of these potential benefits have been realized through the process of building an optimization model.

Underlying Assumptions of the Optimization Model

In building this optimization model, the structures and relationships of other models were analyzed and their underlying assumptions challenged. The optimization model makes different assumptions and, most importantly, the model defines the relationships between outcomes and inputs differently.

The fundamental assumption regarding education is that it is stochastic in nature because the goals of education are mostly measured by student achievement tests having theoretical and practical upward limits. The critical step in actually building the optimization model was identifying the mathematical function fitting the stochastic nature of education to a diminishing returns curve rather than a constant returns line. Considerable attention was paid to the mathematical evidence demonstrating the existence of a diminishing returns curve derived from a transformation of the regression analysis. Using the principles of mathematical programming, it was possible to: (1) Incorporate these diminishing returns curves into multiple regression equations representing the simultaneous educational goals; (2) Incorporate additional equations reflecting the constraints on the organization, most importantly, cost; and (3) Develop the methodology for finding feasible solutions to this optimization model. The optimization model is more sophisticated than other models because these concepts are incorporated; and because they are incorporated, the optimization model more accurately represents the actual situation.

Observations Regarding the Optimization Model

The generalized results of the optimization model suggest different courses of action challenging conventional thinking in several ways. First, there is a unique resource allocation strategy for every school, depending on its starting conditions, rather than a common strategy applying to all schools as is the case with other models. Second, while additional resources can make some difference, merely adding educational resources will never completely overcome the influence of SES or the shortcomings in organizational effectiveness. This distinguishes the optimization model from those that resources can overcome all other shortcomings. Third, in some cases, more is better, but in other cases more (e.g., money) produces little or no increased in benefits. In other models, more is always better. Unquestionably, these findings are in direct contrast to the conventional and somewhat “seat-of-the-pants” thinking prevalent in education today.

Identifying and Testing Potentially More Successful Options

There are many “ifs” in model building. In this case, there is the question of whether the stochastic model presented here has greater logical and mathematical merit than other models. Next is the question of the accuracy of magnitude of the relationships presented. Are the estimates of the influence of SES, school effectiveness, and school inputs reasonable? Assuming the responses to these questions are in the affirmative, then there is the inescapable question: Why focus so much attention on the allocation of school resources when the largest impact on student achievement will come through improving school effectiveness and addressing the issues associated with community SES?

SES poses its own set of problems. First, SES is not a changeable “thing,” at least changed in a way that relates to student achievement. SES is a concept, and researchers employ proxies to measure the concept. The measure usually includes, for example, income, education levels, and verbal aptitude of the mother. No one seriously proposes policy changes in these variables in order to improve student achievement. More likely, the concept of SES represents a set of behaviors associated with families and communities where students test favorably. Is it the amount of time devoted to reading or homework, or the amount of time not devoted to television? Is it the amount of time parents spend talking with their children about school or the amount of time a family engages in serious discussion about the importance of an education? We do not know. It does seem potentially rewarding, however, to find out more about these behaviors and then devise programs for schools, communities, religious organizations, and social service agencies to become more engaged in an way that is likely to bring more success.

Education is not well-suited for testing the optimization model—or any model—through experimentation. State laws, professional attitudes and traditions, and public opinion make it all but impossible to adopt the conclusions of the optimization model into practice. Some expectations of change have been placed on charter schools, but the evidence is not hopeful. Perhaps the critical question is whether using a different model—an optimization model—can have an impact on lawmakers’ actions, professional attitudes, and public opinion?

The Optimization Model as a Paradigm

This article was heavily influenced by Kuhn’s ideas and, especially, his thoughts regarding a “paradigm shift” in *The Structure of Scientific Revolutions*.¹³ The optimization model in the context of a paradigm has a larger purpose: To put all the individual pieces of an educational organization into a single, comprehensive, and logical framework, much like particle physics and the “Standard Model.” With such a framework in place, it is possible to make more sophisticated inquiries and predictions. The results then become the empirical basis for policy decisions. The driving force for a new model was the anomaly presented by regression analysis; that is, regression could not accommodate all the elements and outcomes of the organization simultaneously, and it could not comprehensively respond to the best use of resources questions.

The intent of the optimization model as a paradigm is to demonstrate its greater robustness compared to its competitors in that it substantially adds scope and precision to the “what if” questions. In addition, the model establishes a framework for future research. First, it builds upon the idea of the production function by adding the element of effectiveness with a theoretical basis and

a practical method for its measurement. Second, it incorporates a reformulation of the regression statistics into a type of glue serving to hold the multiple outcomes together with the multiple elements in a comprehensive and mathematically logical way. Finally, it incorporates a mathematical programming methodology for modeling the intricacies of the educational organization.

What is missing? There seem to be at least three major pieces missing for a concerted research strategy: (1) A conceptual structure guiding research efforts; (2) a set of reliable and replicated measurements of the structure elements and their relationship with outcomes; and (3) methods to address technical shortcomings.

Other sciences have conceptual structures guiding research efforts. While there are many illustrations, the periodic table from chemistry serves as an instructive analogy. The periodic table identifies the basic chemical elements by their measurable characteristics. Based on these characteristics, research is directed toward understanding how they interact with one another in more complex situations. What if there were a comparable conceptual structure for educational organizations? What if there were a consensus regarding the structure and elements of the educational organization along the lines presented herein? It would encourage the direct comparison of research results—a type of unification. Like chemistry, additional elements could be included as their unique characteristics and contributions are identified and measured. With a consensus of the structure and elements of an organization, research would focus on what is in common among organizations so the anomalies could be identified and addressed.

What if there were a comprehensive set of measurements estimating the characteristics of these elements and their relationship to outcomes? While they would not be exact, as they are in chemistry, they would fall within ranges, and these ranges would be valuable in seeding the optimization model. While they will undoubtedly be difference estimates, there is no reason to believe the underlying effect of staffing quality or staffing quantity would be different due to the school district or state of residence of a student. Most likely, it is the unique combination of factors making the difference. Therefore, the key is to identify those underlying factors, their magnitudes, and their relationships.

What if there was a concerted effort to address some of the technical shortcomings of this and other models—the multicollinearity among variables, for example? For example, it may be possible to incorporate the multicollinearity into the optimization model by adding defining equations.

Walberg worked on developing a comprehensive framework for the analysis of productivity starting in 1975.¹⁴ (While he developed a method of measuring relationships between outcomes and school variables—effect size—he neither proposed an economic adjustment nor an optimization method.) Levin addressed the important relationship of cost-effectiveness with educational policy,¹⁵ and Monk described the pro’s and con’s of the production function.¹⁶ The optimization model builds on Walberg’s plea for a comprehensive framework, Levin’s push for cost-effectiveness, and Monk’s call for greater sophistication in the production function.

With these caveats in mind, the ultimate value of this model is its potential for becoming a paradigm for the continued pursuit of educational productivity.

Endnotes

¹ Linus Schrage, *Lindo: An optimization Modeling System* (San Francisco: Scientific Press, 1991).

² Thomas S. Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1970).

³ H. P. Williams, *Model Building in Mathematical Programming* (New York: Wiley & Sons, 1985).

⁴ Ibid., 3.

⁵ There are also criterion referenced assessments with the purpose of identifying those students who have achieved a minimum academic standard. The distribution of these assessments is not normal, but more of a "J-shape" with a tail at the lower end and a large grouping at the high end. However, when the assessment scores are grouped by building, they tend to take on the shape of the normal curve.

⁶ For a thoughtful and comprehensive discussion of the production function, see David H. Monk, (New York: McGraw-Hill: *Educational Finance: An Economic Approach*, 1990).

⁷ James L. Phelps and Michael F. Addonizio, "How Much Do Schools and Districts Matter? A Production Function Approach to School Accountability," *Educational Considerations* 33 (Spring 2006): 51-62. .

⁸ Ibid.

⁹ For those who counter with the Glass and Smith meta-analysis on class-size, see Appendix A.

¹⁰ J. P. Gilford, *Fundamental Statistics in Psychology and Education* (New York: McGraw-Hill, 1965), 399.

¹¹ In Excel, the function is NORMDIST.

¹² Gilford, 394-400.

¹³ It maybe helpful to compare Figure 6 with Figure 1.

¹⁴ Schrage, 8.

¹⁵ Combining several variables into an SES index would make the model structure more straightforward.

¹⁶ Note that the school variables were highly correlated with the SES variable.

¹⁷ Kuhn, *The Structure of Scientific Revolutions*.

¹⁸ Herbert J. Walberg, "Improving the Productivity of America's Schools," *Educational Leadership* 41 (May 1984):19-27.

¹⁹ Henry M. Levin, "Cost-Effectiveness and Educational Policy," *Educational Evaluation and Policy Analysis* 10 (Spring (1988): 51-69.

²⁰ Monk, *Educational Finance*.

Appendix A

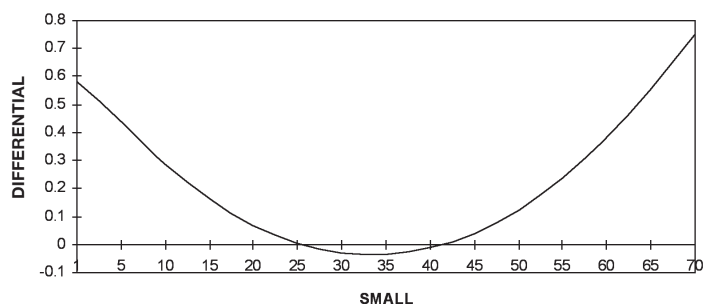
Observations Regarding Meta-Analysis of Class-Size

For those who might cite the class-size meta-analysis by Glass and Smith as an example of increased returns to scale rather than diminishing returns, they may wish to consider the following.¹ First, the equation Glass and Smith used to plot the frequently cited curve included a squared term, indicating the plot is a parabola. When fully plotted across the entire class-size range in the data, the achievement prediction for a class-size of 60 was the same as for a class-size of 10, with the minimum being a class-size of about 32. Because the data included substantial observations of class-size above 40, the full curve should be considered when drawing conclusions rather than just the "attractive" side of the curve. Second, because the report included the data, a re-analysis is possible. When this author conducted a re-analysis, no relationship was found between class-size and achievement levels when the range was restricted to class-sizes between 10 and 60. Third, the class-size scale is not equal interval; therefore it would take four times as many teachers to reach a class-size of 10 starting at 40 as it would to reach 20.

When looking at the entire curve, three first-impression questions come to mind: (1) Can it be that a class-size of 65 will produce the same results as a class-size of 1? (2) What will be the results if there were more teachers than students in the class- would achievement continue to improve? (3) At what class-size does the left-hand side of the curve level off or is perfect achievement attainable? (See Figure A.)

¹ Gene V. Glass, and Mary Less Smith, *Meta-analysis of Research on the Relationship of Class-size and Achievement* (Portland, OR: Far West Laboratory of Educational Research and Development, 1978).

Figure A
Range of Weightings



Appendix B

What Makes Education Stochastic?

After describing much of the details of the stochastic model, it may be useful to revisit the reasons why education evaluation is stochastic. Student achievement tests are based on the properties of the normal or probability curve and administered to students usually during the same grade in school producing another normal-like distribution. This is unlike most outcome measures in other organizations. Therefore, the relationship between student achievement and independent variables should also be based on these same properties. What are these properties?

First and most importantly, the normal curve is bounded. While the curve actually extends from minus infinity to plus infinity, both arms are asymptotic to the abscissa; that is, while the extreme values may

closely approach the boundary, they never do. If in a mathematical model the boundaries could be reached, there would be the “out of bounds” paradox. In the case of education, it would mean all students can be above average, and under some circumstances all students can be perfect. Because this is not the case in practice or in theory, modeling education with stochastic functions more appropriately resembles reality. Second, the relationship among normally distributed variables is nonlinear, a critical condition for solving simultaneous equations. Third, when the predicted results are presented in terms of percentiles, one may answer the question: What are the chances the result will be achieved when the conditions of the model have been met? As the following illustration will show, the changes are limited largely because of the SES element and, to a lesser degree, school effectiveness. In contrast, the regression model implies a 100% chance of achieving perfection given enough resources, regardless of SES or effectiveness.

Because of the stochastic nature of student achievement testing, there is a fundamental difference in how schools are judged compared to most other organizations. All widget-making companies are thought to be successful as long as they stay in business; there is no stochastic judging scheme. While there have been other attempts to judge the performance of schools—for example through accreditation—with the current emphasis on standardized testing, schools have been relegated to a unique fate prescribed by the normal curve.

Teacher Compensation and School Quality: New Findings from National and International Data

**Zhijuan Zhang, Deborah A. Verstegen,
and Hoe Ryoung Kim**

Introduction

High quality education is critical to both the individual and the nation. At the country level, as Ireland's minister for education and science, put it, "The never ending search for competitive advantage in the global knowledge economy has led all public policymakers to focus on education as a key factor in strengthening competitiveness, employment and social cohesion."¹ At the individual level, a student's cognitive achievement is a good predictor of his or her future earnings.²

Compelling evidence shows that the quality of education a school offers influences student achievement.³ Among all variables, teacher quality is the single most important school-related factor affecting student academic achievement.⁴ Teacher quality is at least as important, if not more so, than the socioeconomic status of student family in influencing student academic attainment.⁵ How teachers perform in their classrooms can counteract the negative effects of social, cultural, or human capital.⁶

However, education is challenged by high teacher turnover rates.⁷ The most recent data project that among the 2.2 million new teachers, 666,000 (30%) will leave sometime during their first three years of teaching, and one million (45%) will turn over within the first five years of their teaching career. Teacher turnover is especially problematic in math and science and in many small, high-poverty rural schools.⁸ High teacher turnover rates affect both teacher quantity and quality. When facing a teacher shortage, many school districts either hire underqualified teachers or assign teachers to teach out-of-field. This erodes teacher quality.

Teacher turnover also touches upon issues of social justice and fairness. While research shows that teacher quality matters particularly for students with special needs, low income, low achieving, and minority students are most susceptible to being left in the hands of teachers with lesser skills and knowledge of teaching.⁹ Teachers of these students are more likely to leave when they have obtained some teaching experience.¹⁰ Although out-of-field teaching is widespread, classes in high poverty schools are 77% more likely to be taught by

an out-of-field teacher and staffed with more inexperienced teachers than classes in low poverty schools.¹¹

Around the world, teacher salaries are an important indicator of national or state education priorities and investment. Between 64% and 80% of funding invested in public education is used for paying educational personnel in the OECD¹² countries and in the United States, respectively.¹³ In 2002 alone, the United States invested \$192 billion in teacher pay and benefits.¹⁴ Yet only a few national and fewer international studies have addressed the relationship between teacher salaries and school quality in terms of teacher retention and student achievement. Among them, mixed findings have been found in the U.S. studies,¹⁵ and no evidence has been found supporting a clear relationship across countries between teacher salaries and student achievement.¹⁶ In addition, fewer national and international studies have addressed the relationship between teacher salaries and teacher retention. More often than not, these studies use data for only one specific U.S. state or city limiting generalizability.¹⁷

Are teacher salaries related to school quality in terms of student academic achievement and teacher retention? Are teacher salaries important factors influencing teacher job satisfaction? Is teacher job satisfaction related to retention? This research addressed these questions using international and national data. First, the literature will be briefly reviewed, and then the method and findings will be presented. The final section includes a discussion and implications of the research for practice.

Review of Related Literature

Teacher Salaries and Student Academic Achievement

Among the limited number of studies pertaining to the direct relationship between teacher salary and student academic achievement, mixed findings have been produced. In an examination of extant studies, Hanushek, writing on whether money matters in education—either as a function of teacher salaries, pupil-teacher ratio, equipment or facilities—found it did not.¹⁸ Verstegen and King, examining only those studies with statistically significant findings, found a statistically significant and positive association between teacher salaries and student achievement.¹⁹ They noted that Hanushek reached his conclusions by counting both statistically significant and insignificant studies, a method not endorsed by most researchers. Loeb and Page found a strong impact of teacher salary on teacher quality and argued that "even if school districts are unable to identify teacher quality, one would expect the supply of high-ability teachers to increase with teacher wages."²⁰ They found that previous research did not control for alternative labor market opportunities and non-pecuniary school district characteristics, and resulted in mixed findings.

Despite their limited number, some international studies do address the relationship between the two. For example, Barro and Lee, taking advantage of newly constructed panel datasets which included educational inputs and outputs from a broad number of countries, found that the average salary of primary school teachers has a positive and significant relationship with test scores.²¹ However, most international studies pertaining to the relationship between teacher salaries and student academic achievement have found no clear positive link between teacher salaries and student achievement.²²

Teacher Salaries, Teacher Job Satisfaction, and Teacher Retention

Much of the previous research on teacher retention, whether applying a national or an international model, shares the misassumption

Zhijuan Zhang is an instructor in the Virginia Beach City Public Schools. Deborah A. Verstegen is a Professor in the Department of Educational Leadership at the University of Nevada, Reno. Hoe Ryoung Kim is Director in the Department of Educational Finance at the Seoul Kangdong Office of Education, South Korea.

that “the attrition rate of the existing stock of teachers is insensitive to salaries, and does not vary across subject areas, across regions, or over time.”²³ Following this logic, classic job satisfaction theories emphasize non-pecuniary versus pecuniary rewards as does early research in the field. For example, Choy and her colleagues stated that very few people enter the teaching profession for external rewards such as salary, benefits, or prestige.²⁴ Lortie noted that the teaching profession has long been regarded as having a halo of moral commitment and further observed that the culture of the teaching profession and the structure of rewards de-emphasize extrinsic rewards and encourage intrinsic rewards.²⁵ Sergiovanni²⁶ and Dinham and Scott²⁷ found that teacher salary is a hygiene factor, a factor that only prevents job dissatisfaction but does not generate job satisfaction.²⁸

Moreover, only a small proportion of teacher turnover is found to relate to teacher job satisfaction, which Ostroff attributed to the fact that most former studies were analyzed at the individual level while turnover is more a phenomenon of an organization.²⁹ His work showed that teacher job satisfaction has a robust association with retention when data were aggregated at the organizational level. However, whether this finding occurs at higher levels of aggregation is still unknown.

Although the new wave of research has made a breakthrough by concluding that higher salaries are associated with lower teacher attrition, it is still mainly based on cross-sectional data instead of national data, making generalizability difficult. Meanwhile, most of the reported effects of teacher salaries found in the research have been derived from coefficients on salary in turnover analyses.³⁰ Some new research has managed to analyze the relationship between teacher salaries and teacher retention using national longitudinal data and more advanced analytical techniques, such as Shen’s 1997 study and Ingersoll’s 2001 study.³¹ Surprisingly, even using the same data, their findings pertaining to the effect of teacher salaries on teacher retention were dissimilar. For example, Shen found that the annual salary for all teachers and the salary for senior members influenced teacher retention. Conversely, Ingersoll showed that after controlling for administrative support, student discipline, higher levels of faculty decisionmaking influence, and autonomy, teacher salaries became insignificant at the 90% confidence level. Kelly, in a more recent study of teachers in the 1990-1991 Schools and Staffing Survey and the 1991-1992 Teacher Follow-up Survey, found that for the majority of the teaching career, salaries are positively related to teacher retention although the effect is stronger in the early years. This research seeks to clarify these relationships.³²

Methodology

This study addressed the question of whether teacher salaries relate to school quality in terms of teacher retention and student achievement, and, if so, how. It further examined whether teacher job satisfaction is a strong mediator between teacher salaries and teacher retention.

Two data sources were used for the analysis. The first one was the longitudinal national dataset from the 1999-2000 School and Staffing Survey (SASS) and the 2000-2001 Teacher Follow-Up Survey (TFS), sponsored by the National Center for Education Statistics (NCES). The SASS is the largest national dataset pertaining to teachers, administrators, and the general conditions of American elementary and secondary schools. The TFS has become an inseparable part of SASS: Teachers that responded to the SASS are followed and surveyed a year after each administration of the SASS. The purpose of the TFS

is to track teachers after the SASS school year, including those who have changed schools, left teaching, or stayed in the same school, i.e. stayers, movers, and leavers, respectively.

The second data source was the Programme for International Student Assessment (PISA), which provides internationally comparable evidence on student academic achievement in the year 2000. The PISA was jointly developed by participating countries and administered to 15-year-old students in schools in OECD countries. Since the PISA survey provides little information on teacher salary and educational expenditures, 2000 salary data were downloaded from the OECD web site.³³

For the purpose of this study, the U.S. population was limited to public school teachers who taught students in grades K-12 in school year 1999-2000. Only teachers who answered both the SASS and TFS and stayed at their schools were included in the analysis. The sample size for the dataset was 2,894. We hypothesized that teacher salary is associated with teacher general job satisfaction, which results in teacher retention, an important measure of school quality or school effectiveness. Because the literature suggests that school climate, school poverty, and teacher professional growth also affect teacher job satisfaction, they were entered into the model.

Twenty-eight OECD countries and four non-OECD countries participated in the 2000 PISA assessment. The sample size was 26 countries,³⁴ with Luxembourg and Poland deleted from the analysis due to lack of data and the small sample size. The mathematic scores of students from the OECD were obtained from the PISA dataset by teacher and then aggregated at the country level. The teacher salary variable was measured by the ratio of national average teacher salary after 15 years of experience to the national average teacher starting salary in 2000. Salaries for any position of 20 hours or more per week were included, as were any bonuses. We hypothesized that this ratio has substantial influence on student academic achievement. Teacher salaries were converted to equivalent U.S. dollars and adjusted using Purchasing Power Parities.³⁵

The data analysis procedure was divided into two stages: (1) structural equation modeling analysis of SASS data at a national level; and (2) regression analysis of PISA and its supplementary teacher salary data at an international level.

Analysis and Findings

U.S. Individual Teacher Analysis³⁶

In the first stage, data were weighted by TFS final weights as suggested by NCES to ensure sampled teachers are representative of the K-12 public population. A preliminary analysis was conducted to determine the measurement model, which focused mainly on the relationship between latent variables and their indicators by factor analyzing all the items measuring the same latent variables. SPSS statistical software was used for this analysis. Variables that had double loadings on various factors and that had low commonalities on all factors were deleted.

The baseline model was trimmed based on the results of the factor analysis to include:

(1) school climate, as measured by teacher autonomy, teacher participation in decision making, student school conduct, principal leadership, teacher collegiality, and class attendance;

(2) professional growth, as measured by professional development in content teaching, professional development in performance standards, professional development in teaching method, professional

Table 1
Correlation Matrix

	x1	x2	x3	x4	x5	x6	x7	y1	y2	y3	y4	y5	y6	y7	y8	y9	y10
x1	1.000																
x2	.335	1.000															
x3	.048	.180	1.000														
x4	.260	.457	.318	1.000													
x5	.114	.307	.365	.589	1.000												
x6	.083	.200	.594	.282	.282	1.000											
x7	-.100	-.037	-.071	-.008	.017	-.155	1.000										
y1	.138	.159	.138	.179	.070	.116	-.151	1.000									
y2	.031	.110	.075	.114	.174	.040	.060	-.053	1.000								
y3	-.026	.062	.030	.058	.060	.024	.060	-.117	.389	1.000							
y4	-.008	.081	.046	.033	.053	.000	.060	-.014	.201	.213	1.000						
y5	-.034	.062	-.011	.032	.064	-.043	.108	-.063	.224	.273	.252	1.000					
y6	.029	.143	.046	.127	.106	.092	.065	-.027	.133	.109	.209	.116	1.000				
y7	.187	.237	.265	.309	.222	.200	-.015	.161	.096	.053	.030	.037	.016	1.000			
y8	.186	.222	.152	.202	.140	.118	-.032	.252	.037	.050	.042	.006	-.006	.367	1.000		
y9	.080	.123	.099	.121	.081	.048	.000	.122	.077	.086	.017	.011	.022	.194	.373	1.000	
y10	.020	-.006	.021	-.012	-.004	.045	-.008	.026	.052	.063	.042	.030	-.005	.032	.062	.135	1.000

Where: x1= teacher autonomy; x2=teacher participation in decision making; x3=student behavior; x4=principal leadership; x5=teacher collegiality; x6= school discipline; x7= school poverty; y1=perception of teacher compensation; y2 = professional development in contents; y3=professional development in standards; y4= professional development in methods; y5= professional development in student Assessment; y6= professional development in discipline; y7= feel it a waste of time to try to do one's best as a teacher; y8= will or will not to become a teacher if one can start over again; y9= the length one plans to remain in teaching; y10=teacher retention.

development in student assessment, and professional development in student behavior;

(3) Teacher job satisfaction, as measured by asking whether a teacher regards teaching as a waste of time, whether one would become a teacher again if he or she had an opportunity to start over, and the length one plans to remain in teaching;

(4) teacher salary;

(5) school poverty;

(6) teacher retention.³⁷

All Cronbach coefficients were found to be over .700, indicating very good reliability. One change suggested by the modification index and factor loadings was that teacher autonomy was not a school climate indicator. Regarding its importance in teacher job satisfaction literature, it was retained in the model as a latent factor independent of school climate. Correlation coefficients of the indicators are listed in Table 1. After modifying the baseline model, adequate model fit was achieved:

$\Delta X^2=854.194$, $\Delta df=1$, $p<.05$;

GFI (goodness of fit index)=.964;

AGFI (adjusted goodness of fit index) =.943;

CFI (comparative fit index)=.892;

RMSEA (root mean square error of approximation) =.056.

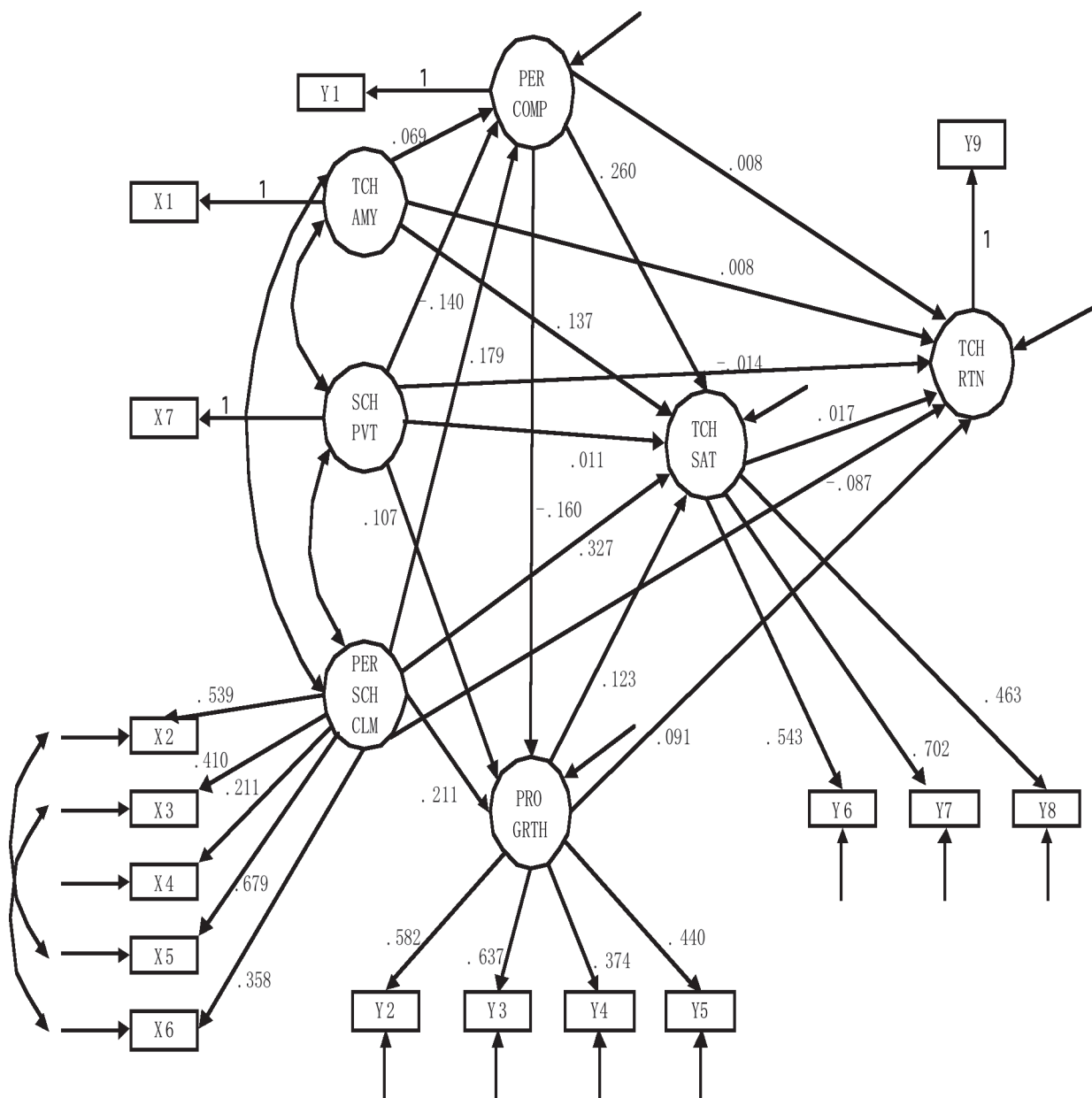
Moreover, all parameter estimates and standard errors were found to be reasonable. Figure 1 shows the streamlined model and the influence of the factors on teacher job satisfaction and retention.

The results showed that approximately 28.6% of the variance of teacher job satisfaction and 2% of the variance of teacher retention was explained by the model. School climate, teacher autonomy, teacher salary, and professional growth had direct and positive effects on teacher job satisfaction. Teacher salary was the second best predictor of teacher job satisfaction with a standardized direct effect of .260, next only to the effect of school climate which was .327. This means that each time when teacher salary goes up by 1, teacher job satisfaction increases by .260 in the model. As related to teacher retention, teacher job satisfaction was found to be the best predictor with a standardized direct effect of .134 in the model. However, no direct association was found between teacher salary and teacher retention.

The path from teacher salary to teacher job satisfaction was further examined by using multigroup analysis to see whether the effect would be impacted by teacher gender, age, years of teaching experience, highest educational degree, and main teaching field. Moreover, some contextual factors suggested by the literature such as school level (elementary or secondary), school size (big or small), and school locality (urban or rural), were also examined.³⁸

No differences in the influence of teacher salary on teacher job satisfaction were found across teachers with differences in length of teaching experience, highest educational degree, or main teaching

Figure 1
Job Satisfaction and Retention Model with Data (Without Movers)



Where: TCH AMY = Teacher Autonomy; SCH PVT = School Poverty; PER SCH CLM = Perception of School Climate; PRO GRTH = Professional Growth; PER COMP = Perception of Compensation; TCH SAT = Teacher Job Satisfaction; TCH RTN = Teacher Retention; X1=Teacher Autonomy; X2=Teacher Participation in Decision Making; X3=Student behavior; X4=Principal Leadership; X5=Teacher Collegiality; X6= Class Attendance; X7= School poverty; Y1=Perception of Teacher Compensation; Y2 = Professional Development in Contents; Y3 =Professional Development in Standards; Y4= Professional Development in Methods; Y5= Professional Development in Student Assessment; Y6= Feel it a waste of time to try to do one's best as a teacher; Y7=Will or not to become a teacher if one can start over again; Y8= The length one plans to remain in teaching; Y9=Teacher Retention.

fields. No differences were found across teachers in schools of different levels, sizes, or locations. However, paths from teacher salaries to teacher job satisfaction were found not to be equivalent across teachers at different ages and with different lengths of teaching experience. The path is equivalent across the group of teachers with over 5 years but less than 20 years teaching experience and the group of teachers with over 20 years teaching experience. Therefore, these two groups were combined into one group, namely, teachers with over 5 years teaching experience. Although the finding that teacher salaries were good predictors of teacher job satisfaction remained robust, the degree of association between teacher salaries and teacher job satisfaction differed across the group of teachers with 5 years or less teaching experience and the group of teachers with more than 5 years teaching experience. As shown in Table 2, compared to teachers with over 5 years teaching experience, teachers with 5 years or less teaching experience were less likely to be dissatisfied by low teacher salaries.

Also the data showed that the association between teacher salaries and teacher job satisfaction was significant across all age groups, but the degree of association differed across teachers less than 50 years old and teachers of 50 years or more. (See Table 3.) Although for all teachers, teacher salary was significantly associated with job satisfaction, the association was less strong for teachers 50 years and over. For these teachers, every change in teacher salary was only associated with a change of .091 in teacher job satisfaction while the association between these two variables for the other two groups was .138. This means that, compared to other teachers, teacher salary was less important to the job satisfaction of teachers 50 and over.

Based on the research results, a post-hoc analysis was conducted. Together with teacher salary, teacher participation in decisionmaking, principal leadership, student discipline, student preparedness to learn, and teacher collegiality were entered in the model. Teacher salary and each of the school climate factors were hypothesized to directly affect teacher job satisfaction and teacher retention.

The model fit the data adequately:

$\Delta X^2 = 537$, $\Delta df = 21$, $p < .05$;

GFI (goodness of fit index) = .935;

AGFI (adjusted goodness of fit index) = .918;

CFI (comparative fit index) = .909;

RMSEA (root mean square error of approximation) = .052.

The results are presented in Figure 2. Findings showed that teacher salaries and teacher participation in decisionmaking were the two most important determinants of teacher job satisfaction. The difference between them was 0.003, which is insignificant.

OECD Analysis³⁹

International data from OECD countries including teacher salary data were analyzed at this stage to determine the relationship between teacher salary and student achievement. Descriptive statistics for the independent variables and dependent variable are presented in Table 4. Canada, Netherlands, and New Zealand had some missing data, and these descriptive statistics were computed by list-wise deletion. Table 4 shows a large range between minimum teacher salary and maximum teacher salary, and between minimum expenditure on lower secondary education per student and maximum expenditure on lower secondary education per student. For example, maximum teacher salary was about seven times greater than minimum salary in both starting teacher salary and teacher salary after 15 years of experience. Maximum

Table 2

Group Comparison of Effects of Teacher Salary on Teacher Job Satisfaction Based on Length of Teaching (in Years)

Group Comparison	Between Group Differences	
Group 1 Group 2	No	
Group 2 Group 3	Yes	b = .122*
		b = .133*
Group 1 Group 3	Yes	b = .097*
		b = .133*

*P < .05

Where: Length of teaching experience for Group 1 > 5 years; Group 2 > 5 years and < 20 years; and Group 3 > 20 years.

Table 3

Group Comparison of Effects of Teacher Salary on Teacher Job Satisfaction Based on Age

Group Comparison	Between Group Differences	
Group 1 Group 2	No	
Group 2 Group 3	Yes	b = .139*
		b = .093*
Group 1 Group 3	Yes	b = .138*
		b = .093*

*P < .05

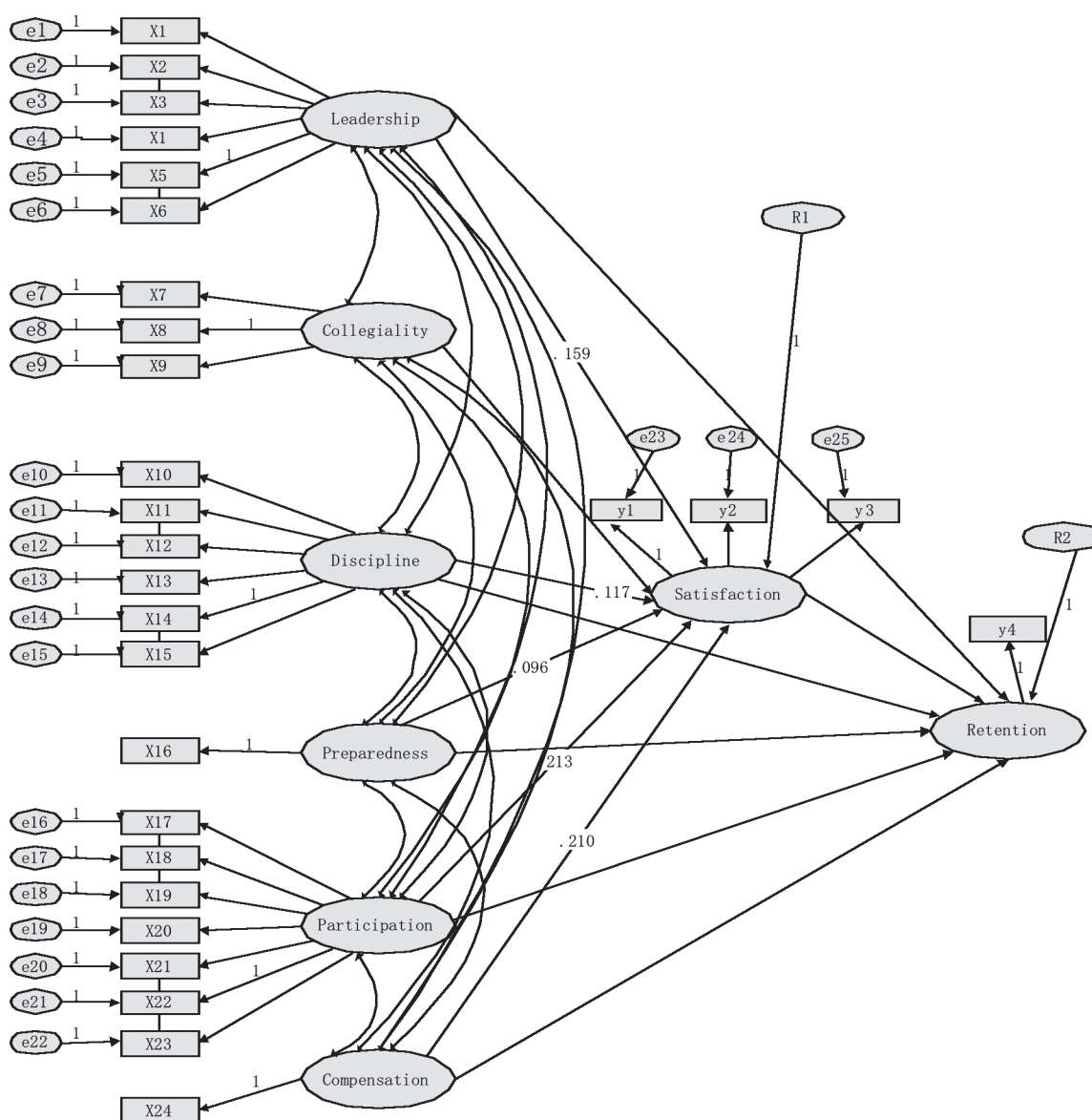
Where: Age for Group 1 < 40 years; Group 2 > 40 and ≤ 50 years; and Group 3 > 50 years.

educational expenditure per student was also about seven times as much as minimum educational expenditures per student across 26 OECD member countries.

Correlation coefficients presented in Table 5 indicate that national average math test scores were highly correlated with the ratio of teacher salary after 15 years of experience to teacher starting salary ($r = .450$; $p \leq 0.05$). Moreover, it also showed that national average math test scores were more strongly related to teacher salary after 15 years of experience ($r = .438$; $p \leq 0.05$) than teacher starting salary ($r = .224$; $p \leq 0.05$). As in the United States, teacher salary is a major portion of expenditure per student in the OECD countries, and Table 5 also shows that there was a strong correlation between expenditure per student on lower secondary education and teacher starting salary ($r = .598$; $p \leq 0.05$) and teacher salary after 15 years of experience ($r = .520$; $p \leq 0.05$).

Table 6 presents the results of a regression model where the dependent variable was mean national math test scores and the

Figure 2
Post-hoc School Climate and Compensation Model



Where: Leadership=Principal Leadership; Collegiality=Teacher Collegiality; Discipline=Student Discipline; Preparedness=Student Preparedness To Learn; Participation=Teacher Participation In Decision Making; Compensation= Teacher Perceived Compensation; Satisfaction=Teacher Job Satisfaction; Retention=Teacher Retention.

independent variables were expenditure per student on lower secondary education and the ratio of teacher salary after 15 years of experience to teacher starting salary. The independent variables accounted for about 50% of the variance in national math test scores among the 26 OECD countries. Based on the F-test, regression coefficients were determined to be statistically significant: b_1 : $F_{1,23} = 12.21$, $p \leq 0.05$; b_2 : $F_{1,23} = 11.83$, $p \leq 0.05$.

The results indicated that if everything else were equal, for every one standard deviation unit change in the ratio of teacher salary after 15 years of experience to teacher starting salary, a .548 standard deviation unit change in national mean math test scores in the same direction would be expected. Similarly, if everything else were equal, for every one standard deviation unit change in expenditure per student on lower secondary education, a .539 standard deviation unit

changes in national mean math test scores would be expected in the same direction. Thus, these results suggest that compensating experienced teachers adequately and overall level of per pupil expenditure predicted higher student academic achievement in secondary math across countries.

The unique contribution of each b_1 and b_2 in accounting for the proportion of variance in national mean math test scores was investigated by conducting hierarchical modeling. Hierarchical modeling compares the full regression model with all predictors to a reduced regression model with fewer predictors than the full model. Based on the results of hierarchical modeling, the unique contribution of b_1 and b_2 in accounting for the variance in national mean math test scores was 28.3 % and 21.4 %, respectively. The F-test showed that the unique contributions of b_1 and b_2 were both statistically significant:

Table 4
Descriptive Statistics of Variables in OECD Analysis

Variables	N	Minimum	Maximum	Mean	Standard Deviation
National teacher starting salary	25	6,340	41,358	23,980.32	7,732.72
National teacher salary after 15 years of experience	25	8,957	54,852	32,722.42	10,339.84
Ratio of teacher salary after 15 years of experience to teacher starting salary	25	1.11	1.93	1.37	22.02
Expenditure on lower secondary education per student	25	1,289	8,934	5,877.60	1,941.60
National average math test scores	25	387	557	503.32	37.38

Table 5
Correlation Matrix of Variables in the OECD Analysis

Variables	Starting Salary	Salary after 15 Years of Experience	Country Mean Math Scores	Expenditure Per Student on Lower Secondary Education	Ratio of Salary after 15 Years of Experience to Starting Salary
Starting Salary	1.000				
Salary after 15 Years of Experience	.882**	1.000			
Country Mean Math Scores	.224	.438*	1.000		
Expenditure on Lower Secondary Education Per Student	.598**	.520**	.462*	1.000	
Ratio of Salary after 15 Years of Experience to Starting Salary	-.209	.267	.450*	-.161	1.000

*P ≤ .05.

**P ≤ .01.

Table 6
Regression Coefficients

	Unstandardized Coefficients		Standardized Coefficients	t	Significance
	B	Std. Error	Beta		
Constant	315.120	43.589		7.229	.000
b_1	.011	.003	.548	3.494	.002
b_2	.915	.266	.539	3.439	.002

Where: Dependent Variable=country mean math scores; b_1 =Ratio of teacher salary after 15 years of experience to teacher starting salary; b_2 =Expenditure per student on lower secondary education.

b_1 : $F_{1,23} = 5.983$, $p \leq 0.05$; b_2 : $F_{1,23} = 5.983$, $p \leq 0.05$.

Discussion and Implications

Teacher job satisfaction was found to be a good predictor of teacher retention, and among all the factors that directly relate to teacher job satisfaction in the streamlined model, teacher salary was the second most important, only next to school climate. A better school climate was found to be associated with greater teacher job satisfaction. In

addition, the indicators of school climate, including teacher participation in decisionmaking, student school conduct, principal leadership, teacher collegiality, and class attendance, all positively contributed to a good school climate that elicited greater teacher job satisfaction and potentially increased teacher retention rates.

In the final post hoc analysis examining the importance of teacher salary, teacher salary stood out as important as teacher participation in decisionmaking in predicting teacher job satisfaction, and,

consequently, teacher retention. Moreover, the results of the multigroup analyses showed that teacher salary was a strong predictor of teacher job satisfaction despite teacher age, length of teaching experience, gender, major field of teaching, or highest educational degree earned, and despite the level, size, and location of the school where he or she taught. Nevertheless, the multigroup national analysis based on teacher age and the length of teaching experience suggested that the association between teacher salary and job satisfaction and, in turn, teacher retention, was stronger among some teachers. For example, novice teachers who had taught 5 years or less and teachers 50 and over were less concerned about salary than those in other groups.

The results of the international analysis indicated that teacher salary was associated with secondary math test scores along with school resources such as class size, student-teacher ratio, teacher major, quality of instructional resources, and teacher morale. The educational expenditure per student on lower secondary education and the ratio of teacher salary after 15 years of experience to starting salary (salary ratio) together accounted for about 50% of the variance in student academic achievement, which was measured by national average math test scores among 26 OECD member countries. In particular, the salary ratio explained more of the proportion of the variance (28.3%) in student academic achievement among countries than did educational expenditures per student (21.4%). This finding converged with the result of our first stage analysis that money matters, but how effectively educational money is invested and deployed is also important in producing desirable school quality as measured by teacher retention and student academic achievement.

In sum, the findings from this study in the national level analysis confirmed the current research that teacher quality is crucial in student academic achievement.⁴⁰ Thus, ensuring a highly-qualified teaching force for all students should be a national priority in educational policies related to student academic achievement. Increasing current teacher salaries and providing participatory decisionmaking are two key factors in reaching this goal. Furthermore, the international findings from this study indicated that those countries with a steeper salary schedule, have higher national math test scores. Larger and continuing increases in salaries over a teacher's career should be considered by policymakers. The findings from this study supported the importance of both higher teacher compensation and reform in the structure of teacher compensation.

Endnotes

¹ Noel Dempsey, "Building the Knowledge Society," *OECD Observer* 242 (March 2004): 7-8. See also George Psacharopoulos, "The Value of Investment in Education: Theory, Evidence, and Policy," *Journal of Education Finance* 32 (Fall 2006): 113-132.

² David Card and Alan Krueger, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy* 100 (February 1992): 1-40; David Card and Alan Krueger, "School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina," *Journal of Economic Perspectives* 10 (August 1996): 31-50; Eric A. Hanushek, "The Economics of School Quality," *German Economic Review* 6 (August 2005): 269-286; Stephen P. Heyneman, "International Education Quality," *Economics of Education Review* 23 (August 2004): 441-452; Alan C. Kerckhoff, Stephen W. Raudenbush, and Elizabeth Glennie, "Education, Cognitive Skills, and Labor Force Outcomes," *Sociology of Education* 74 (January 2001): 1-24; Richard J.

Murnane, John B. Willett, and Frank Levy, "The Growing Importance of Academic Skills in Wage Determination," *Review of Economics and Statistics* 77 (May 1995): 251-266.

³ Heyneman, "International Education Quality."

⁴ Linda Darling-Hammond and Laura Post, "Inequality in Teaching and Schooling: Supporting High-Quality Teaching and Leadership in Low-Income Schools," in *A Nation at Risk: Preserving Public Education as an Engine for Social Mobility*, ed. Richard D. Kahlenberg (New York: Century Foundation, 2000), 127-167; Dan D. Goldhaber, Dominic J. Brewer, and Deborah J. Anderson, "A Three-Way Error Component Analysis of Educational Productivity," *Education Economics* 7 (December 1999): 199-208; Rob Greenwald, Larry V. Hedges, and Richard D. Laine. "The Effect of School Resources on Student Achievement," *Review of Educational Research* 66 (Fall 1996): 361-396; Larry V. Hedges, Richard D. Laine, and Rob Greenwald, "Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes," *Educational Researcher* 23 (April 1994): 5-14; Jennifer King Rice, *Teacher Quality: Understanding the Effectiveness of Teacher Attributes*, (Washington, DC: Economic Policy Institute, 2003); William L. Sanders and Sandra P. Horn, "Research Findings from the Tennessee Value-Added Assessment System (TVAAS) Database: Implications for Educational Evaluation and Research," *Journal of Personnel Evaluation in Education* 12 (September 1998): 247-256; Deborah A. Verstegen and Richard A. King, "The Relationship Between School Spending and Student Achievement: A Review and Analysis of 35 Years of Production Function Research," *Journal of Education Finance* 24 (Summer 1998): 243-62; Andrew J. Wayne and Peter Youngs, "Teacher Characteristics and Student Achievement Gains: A Review," *Review of Educational Research* 73 (Spring 2003): 89-122.

⁵ Darling-Hammond and Post, "Inequality in Teaching and Schooling"; Steven G. Rivkin, Eric A. Hanushek, and John F. Kain, "Teachers, Schools, and Academic Achievement," *Econometrica* 73 (March 2005): 417-458.

⁶ John Schacter and Yeow Meng Thum, "Paying for High- and Low-Quality Teaching," *Economics of Education Review*, 23 (August 2004): 411-430; Harold Wenglinsky, *How Teaching Matters: Bringing the Classroom Back into Discussions of Teacher Quality* (Santa Monica, CA, and Princeton, NJ: Milken Family Foundation and Educational Testing Service, October 2000); Harold Wenglinsky, *Teaching the Teachers: Different Settings, Different Results* (Princeton, NJ: Educational Testing Service, 2000); Harold Wenglinsky, "Closing the Racial Achievement Gap: The Role of Reforming Instructional Practices," *Education Policy Analysis Archives* 12 (November 2004), <http://epaa.asu.edu/epaa/v12n64>.

⁷ Richard M. Ingersoll, "The Problem of Underqualified Teachers in American Secondary Schools," *Educational Researcher* 28 (March 1999): 26-37; Richard M. Ingersoll, "Teacher Turnover and Teacher Shortage: An Organizational Analysis," *American Educational Research Journal* 38 (Fall 2001): 499-534; Richard M. Ingersoll and Thomas M. Smith, "The Wrong Solution to the Teacher Shortage," *Educational Leadership* 58 (May 2003): 30-45.

⁸ Richard M. Ingersoll, *Is There Really a Teacher Shortage?* (Seattle, WA: Center for the Study of Teaching and Policy, University of Washington, September (2003); David H. Monk, "Recruiting and Retaining High-Quality Teachers in Rural Areas," *The Future of Children* 17 (Spring 2007): 155-173.

⁹ Darling-Hammond and Post, "Inequality in Teaching and Schooling"; Linda Darling-Hammond and Gary Sykes, "Wanted: A National Teacher Supply Police for Education: The Right Way to Meet the 'Highly Qualified Teacher' Challenge," *Education Policy Analysis Archives* 11 (September 2003), <http://epaa.asu.edu/epaa/v11n33>; Hamilton Lankford, Susanna Loeb, and James Wyckoff, "Teaching Sorting and the Plight of Urban Schools: A Descriptive Analysis," *Educational Evaluation and Policy Analysis* 24 (Spring 2002): 37-62; Deborah A. Verstegen, "Efficiency and Equity in the Provision and Reform of American Schooling," *Journal of Education Finance* 20 (Summer 1994): 107-131.

¹⁰ Darling-Hammond and Post, "Inequality in Teaching and Schooling."

¹¹ Craig D. Jerald and Ricard M. Ingersoll, *All Talk, No Action: Putting an End to Out-of-Field Teaching* (Washington, DC: The Education Trust, August 2002); Heather G. Peske and Kati Haycock, *Teacher Inequality: How Poor and Minority Students Are Shortchanged on Teacher Quality* (Washington, DC: The Education Trust, June 2006).

¹² Organisation for Economic Co-Operation and Development, www.oecd.org.

¹³ Helen F. Ladd, "Teacher Labor Markets in Developed Countries," *The Future of Children* 17 (Spring 2007): 201-217; Thomas B. Parrish, Christine S. Hikido, and William J. Fowler, *Inequalities in Public School District Revenues* (Washington, DC: National Center for Education Statistics, U.S. Department of Education, 1998).

¹⁴ Rice, *Teacher Quality*.

¹⁵ See also, Dale Ballou and Michael Podgursky, *Teacher Pay and Teacher Quality* (Kalamazoo, MI: Upjohn Institute for Employment Research, 1997); Eric A. Hanushek, John F. Kain, and Steven G. Rivkin, "Do Higher Salaries Buy Better Teachers?" A paper presented the annual meeting of the American Economic Association, New York, January 1999; Susanna Loeb and Marianne E. Page, "Examining the Link Between Teacher Wages and Student Outcomes: The Importance of Alternative Labor Market Opportunities and Non-Pecuniary Variation," *Review of Economics and Statistics* 82 (August 2000): 393-408; Verstegen and King, "The Relationship between School Spending and Student Achievement."

¹⁶ Ladd, "Teacher Labor Markets in Developed Countries."

¹⁷ Cassandra M. Guarino, Lucrecia Santibanez, and Glenn A. Daley, "Teacher Recruitment and Retention: A Review of the Recent Empirical Literature," *Review of Educational Research* 76 (Summer 2006): 173-208; Jonathan H. Mark and Barry D. Anderson, "Teacher Survival Rates in St. Louis, 1969-1982," *American Educational Research Journal* 22 (Fall 1985): 413-421; Richard J. Murnane and Randall J. Olsen, "The Effects of Salaries and Opportunity Costs on Length of Stay in Teaching: Evidence from North Carolina," *The Journal of Human Resources* 25 (Winter 1990): 106-124; Richard J. Murnane, John B. Willett, and Frank Levy, "The Growing Importance of Academic Skills in Wage Determination," *Review of Economics and Statistics* 77 (May 1995): 251-266.

¹⁸ Eric A. Hanushek, "The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economics Literature* 24 (September 1986): 1141-1147.

¹⁹ Verstegen and King, "The Relationship between School Spending and Student Achievement."

²⁰ Loeb and Page, 393.

²¹ Robert J. Barro and Jong-Wha Lee, "Schooling Quality in a Cross-Section of Countries," *Economia* 68 (November 2001): 465-488.

²² Ladd, "Teacher Labor Markets in Developed Countries."

²³ Richard J. Murnane and Randall J. Olsen, "The Effects of Salaries and Opportunity Costs on Duration of Teaching: Evidence From Michigan" *Review of Economics and Statistics* 71 (May 1989): 245.

²⁴ Susan P. Choy, Sharon A. Bobbitt, Robin R. Henke, Elliott A. Medrich, Laura J. Horn, and Joanne Lieberman, *America's Teachers: Profile of a Profession* (Washington, DC: National Center for Education Statistics, U.S. Department of Education, 1993).

²⁵ Dan Clement Lortie, *Schoolteacher: A Sociological Study* (Chicago: University of Chicago Press, 1975).

²⁶ Thomas Sergiovanni, "Factors which Affect Satisfaction and Dissatisfaction of Teachers," *The Journal of Educational Administration* 1 (October 1967): 66-82

²⁷ Steve Dinham and Catherine Scott, "Moving into the Third, Outer Domain of Teacher Satisfaction," *Journal of Educational Administration* 38 (August 2000): 379-396.

²⁸ See Frederick Herzberg, *Work and the Nature of Man* (Cleveland: The World Publishing Company, 1966).

²⁹ Cheri Ostroff, "The Relationship between Satisfaction, Attitudes, and Performance: An Organizational Level Analysis," *Journal of Applied psychology* 77(December 1992): 963-974.

³⁰ Guarino et al., "Teacher Recruitment and Retention."

³¹ Jianping Shen, "Teacher Retention and Attrition in Public Schools: Evidence from SASS91," *The Journal of Educational Research* 91 (November 1997): 81-88; Ingersoll, "Teacher Turnover and Teacher Shortage."

³² Sean Kelly, "An Event History Analysis of Teacher Attrition: Salary, Teacher Tracking, and Socially Disadvantaged Schools," *Journal of Experimental Education*, 72 (March 2004): 195-220.

³³ Organisation for Economic Co-Operation and Development Online Education Database, www.oecd.org/education/database.

³⁴ These countries included: Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Japan, Korea, United States, Mexico, Netherlands, Norway, New Zealand, United Kingdom, Portugal, Greece, Hungary, Iceland, Ireland, Italy, Spain, Sweden, and Switzerland.

³⁵ See, "Purchasing Power Parities (PPP)," http://www.oecd.org/departement/0,3355,en_2649_34357_1_1_1_1_1,00.html.

³⁶ For more details, see Zhijuan Zhang, "Teacher Job satisfaction and Retention: A study of K-12 Teachers Using Structural Equation Modeling" (Ph.D. diss., University of Virginia, 2005)

³⁷ Teacher poverty and teacher retention each has only one indicator, and therefore no measurement error is supposed to exist for both of them.

³⁸ Guarino et al., "Teacher Recruitment and Retention."

³⁹ For more details, see Hoe Ryoung Kim, "The Role of School Resources in Student Achievement in OECD Countries: A Structural Equation Modeling Analysis" (Ph.D. Diss., University of Virginia, 2004).

⁴⁰ Darling-Hammond and Post, "Inequality in Teaching and Schooling"; Rice, *Teacher Quality*.

Explaining the Relationship Between Resources and Student Achievement: A Methodological Comparison of Production Functions and Canonical Analysis¹

Robert C. Knoeppel and James S. Rinehart

What is the relationship between inputs to education and student achievement? The elusive answer to this seemingly self-evident question has led some to characterize the question as the “holy grail” of school finance research for the past thirty years.² Previous attempts to answer this important research question have relied primarily on the use of education production functions. Although the reliance on this method has led to mixed results, the literature base reveals that recent studies have shown a positive, robust relationship between inputs to schooling and measures of student achievement.³ These studies examine not just dollar inputs to schooling, but what those dollars purchased, such as teacher characteristics, class sizes, curriculum, technology, and facilities. Monk notes that one way to combat inconsistent results in production function studies is for researchers to conduct separate studies using different data, methodological designs, and statistical techniques that may confirm previous results.⁴ He postulates that the use of the education production function is flawed because this methodology only relates to education productivity in a marginal way. The use of a single output is an inadequate description of the production relation that may exist in a school given the multiple dimensions of schooling.

Toward that end, an emerging body of literature has begun to examine the relationship between resources for education and measures of student achievement by making use of multiple dependent measures. Schwartz, Stiefel and Hadj made use of cost functions to measure the performance of elementary, middle and high schools in Ohio over a three year period to discern the minimum cost of producing a bundle of outputs given a particular technology and the price of inputs.⁵ Their analysis revealed a positive relationship between input prices and costs but no relationship between school-level pass rates and funding. Similarly, Rubenstein made use of multiple output

variables to assess school efficiency using a methodology entitled data envelopment analysis (DEA).⁶ DEA is a linear programming technique that makes use of a nonparametric efficiency frontier that includes all decision making units in the sample. Using this method of analysis, the researcher found groups of schools that were performing better than would be expected given the composition of their population (efficient schools) that he identified for further research. Although not employed in the extant research, canonical analysis is another methodology that may be used to study the relationship between two sets of variables.⁷

This study compared the results from an education production function with those found using canonical analysis. The purpose of this study was to examine the utility of canonical analysis by policymakers. By examining differing methodologies, conclusions may be drawn with regard to efficiency. Educational efficiency is concerned with the use of scarce resources. It is defined as the amount of knowledge “delivered to” and “acquired by” students given a specific set of resources.⁸

Education Production Functions

Previous attempts to find a relationship between resources and student achievement have relied primarily on education production functions. The production function is a statistical technique that describes the maximum level of outcome possible from different combinations of inputs. The existence of a production function infers that there is something systematic about the transformation of inputs into outcomes.⁹ Previous studies have made use of inputs such as resources, organizational characteristics, and student attributes while outputs have included measures of student achievement. These output measures may take the form of level scores, gain scores, or difference scores.¹⁰ For the purpose of practice, knowledge of the process through which inputs are transformed to educational outputs would assist educational leaders and policymakers to make more accurate assessments of efficiency.

Multiple Regression

An example of a production function that utilizes a statistical technique to analyze the relationship between school resources and student learning is multiple regression analysis. This analysis includes two distinct purposes, *correlation* and *regression*, even though the terms are used interchangeably. First, regression analysis is a technique to find the relationship between one dependent variable and two or more independent variables, which is multiple correlation.¹¹ A second purpose is to predict future outcomes based upon analyzing an outcome measure from several independent variables. Both purposes can be utilized in interpreting the outcomes when multiple regression is used as a technique to analyze production function data.¹²

One use of multiple regression in education is to explain student learning based upon inputs found in school settings.¹³ Cohen and Cohen suggest that as “the number of potential causal factors increase, their representation in measures becomes increasingly uncertain, and weak theories abound and compete.”¹⁴ Thus, explaining student learning is a difficult task, and most of the schooling variables are not well-defined. Nonetheless, one might consider years of teaching experience (EXP), amount of funds spent on instruction (FUNDS) or the number of students on free and reduced lunches (FREE) as inputs to account for the variation in student achievement. In a research design using multiple regression, student achievement (SA) can be the dependent variable (Y) and the independent variables (X_i) are the inputs to account for the variance in Y. Given the variables just mentioned, the

Robert C. Knoeppel is Assistant Professor of Education Leadership Studies at the University of Kentucky. James S. Rinehart is Professor and Chair of the Department of Educational Leadership and Policy Studies at Eastern Kentucky University.

multiple regression equation becomes:

$$Y = a + B_1X_1 + B_2X_2 + B_3X_3$$

or

$$SA = a + B_1EXP_1 + B_2FUNDS_2 + B_3FREE_3.$$

$B_1 \dots B_3$ are regression coefficients, and when they are standardized, the relative explanatory power of the independent variables can be compared.

Another important output from multiple regression analysis is the correlation between the independent variables and the dependent variable, which is known as the squared multiple correlation coefficient (R^2) and indicates the amount of variance in the dependent measure accounted for by the independent variables. Thus, in the case in the preceding paragraph, the amount of variance in student achievement can be estimated from the effects of teaching experience, instructional funding, and number of students on free and reduced lunches.

Although outputs from regression analysis may be important, there are conditions that must be met to interpret the analysis results with some certainty. For example, most authors agree that it is important to have the appropriate cases to independent variables, absence of multicollinearity and singularity, and normality and linearity.¹⁵ Thus, the above conditions must be analyzed before attempting to interpret the regression coefficients and multiple correlation.

Criticism of Production Function Studies

Education production formulas, also known as input-output or cost-quality analyses, were highlighted in the 1966 publication, *Equality of Educational Opportunity*, or the “Coleman” Report. This report attempted to ascertain the amount of inequality in America’s schools. While attempts had been made previously to determine this information, no other studies went into as much depth as the Coleman Report nor did they have as far reaching an impact. Succinctly stated, the Coleman Report found that families, and to a lesser extent peers, are the primary determinants of variations found in student performance rather than educational inputs.¹⁶ These results have been controversial, and some scholars have found methodological flaws in the analysis. Numerous studies have followed to attempt to find more evidence supporting the relationship between inputs to schooling and student achievement with Effective Schools research heralding a shift in thinking only to be followed by several well-designed small scale studies that found positive relations for specific resource inputs e.g. class sizes, quality preschool, and quality teachers.¹⁷

Although the use of education production functions has been prevalent in the research concerning the relationship between resource inputs into schooling and student performance, it has been argued that the use of this method of analysis is limited and that education production functions relate to productivity only in a marginal way.¹⁸ The method of analysis is limited in part because it attempts to link the use of inputs to one measure of output: primarily minimum competency test scores.¹⁹ As such, the use of this method provides a poor estimate of the efficiency with which resources are transformed in to student achievement measures. Further, researchers contend that the use of a single output measure is an inadequate description of the production relation that may exist in a school given the multiple dimensions of schooling and multiple goals and objectives.

Another issue is that the use of the education production function has led to apparently different conclusions using the same set of data.

For example, Hanushek²⁰ and Hedges, Laine and Greenwald²¹ report entirely different conclusions as to the effect of increasing funding for public education from the same set of data. Citing 187 “qualified” studies of both single and multiple districts that made use of education production functions, Hanushek concluded that there is no “systematic” relationship between expenditures and student performance.²² As a result, he finds, educational policy should not be formulated solely on the basis of expenditures. Conversely, Hedges, Laine, and Greenwald reanalyzed the data finding fundamental flaws in the research design used by Hanushek while reaching a decidedly different conclusion.²³ The basic argument is that the method of analysis used by Hanushek, vote counting, is problematic when used as a procedure that would enable a researcher to make inferences and that Hanushek uses both significant and insignificant results to reach conclusions. Instead, Hedges, Laine, and Greenwald made use of two forms of meta-analytic techniques to ascertain the effect on student performance of a change in resources made available to schools. Their findings show strong support for resource inputs on student achievement.

Monk addresses the issue of the lack of systematic evidence from production functions. He notes that one possibility for this finding is that there may actually be multiple education production functions at work.²⁴ Perhaps the transformation of inputs to outputs changes based on gender, ethnicity, or subject taught. As such, regularities in the relationship between inputs to schooling and output measures of schooling will only be found when conditions are “so circumscribed that only unique events are captured.”²⁵

Canonical Analysis

Although not frequently employed in the extant research, another methodology that can accommodate multiple inputs and outputs of schooling that is used in this research, canonical analysis, is designed to study the relationship between two sets of variables.²⁶ Conceptually, canonical analysis and multiple regression are similar in terms of purpose and assumptions. The two methodologies differ in that canonical analysis enables the researcher to include multiple dependent measures. According to Thompson, a multivariate method of analysis can better simulate the reality from which the researcher is making generalizations.²⁷ Because researchers care about multiple outcomes, and because outcomes are the result of myriad factors, the chosen method of analysis must honor the researcher’s view of reality otherwise there will be a distortion of results.²⁸ Canonical analysis is a multivariate method of analysis that subsumes other parametric techniques such as t-tests, analysis of variance, regression, and discriminant analysis.²⁹ This method of analysis prevents the researcher from discarding the variance of any variable and it allows one to portray a more accurate picture of reality.³⁰

In canonical analysis, two linear combinations are formed, one of the predictor variables and one of the criteria variables, by differentially weighting them so that the maximum possible relationship between them is obtained. These linear combinations are referred to as the canonical variates, and the relationship between the canonical variates is called the canonical correlation, R_c . The square of the canonical correlation, R_c^2 , is an estimate of the variance shared by the two canonical variates. It is not an estimate of the variance shared between the predictors and criteria but rather of the linear combination of these variables.³¹

Like multiple regression, canonical analysis seeks a set of weights that will maximize a correlation coefficient. In fact, multiple regression

may be considered to be subsumed under canonical analysis because when using only one dependent variable, canonical analysis is reduced to multiple regression. Unlike multiple regression, in which only the X's are differentially weighted, in canonical analysis both the X's and the Y's are differentially weighted. The formula for the linear combination of independent variables may be written as follows:

$$p = b_1y_1 + b_2y_2 + b_3y_3 + b_4y_4 + b_5y_5 + b_6y_6 + \dots + b_{nyn}$$

where p equals the linear combination of independent variables, b equals the standardized canonical coefficient, and y equals the variable. Similarly, the formula for the linear combination of dependent variables may be written as follows:

$$q = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + \dots + a_nx_n$$

where q equals the linear combination of dependent variables, a_i equals the standardized canonical coefficient and x_i represents each of the dependent variables. Canonical correlation finds the relationship between p and q. After having obtained the maximum R_c in canonical analysis, additional R_c 's are calculated, subject to the restriction that each succeeding pair of canonical variates of the X's and the Y's not be correlated with all the pairs of canonical variates that precede it. Like factor analysis and discriminant analysis, the first canonical correlation will probably not account for all of the variance in the data.³²

In canonical analysis, the canonical correlations are calculated in descending order of magnitude, as in discriminant analysis. The first pair of linear combinations is the one that yields the highest R_c possible in a given data set. The second R_c is based on the linear combinations of predictor and criterion variables that are not correlated with the first pair and that yield the second largest R_c possible in the given data set. The same calculation follows for succeeding R_c 's with the maximum number of R_c 's extracted equal to the number of variables in the smaller set when $p \neq q$. A test of significance exists for each canonical correlation and for the total amount of variance accounted for in the two sets of variables. In addition to more scientific tests of significance, the literature suggests that canonical correlations that explain less than ten percent of the shared variance are considered to be not meaningful.³³

Monk argues that chosen methodologies must accommodate for myriad contingencies.³⁴ Canonical correlation is most likely to be useful in situations where there is doubt that one variable can serve as a suitable criterion variable.³⁵ Therefore, by determining if a set of predictor variables correlates with a set of criterion variables, a clearer picture of the relationship between the X and Y variables may be found. It is for these reasons, that canonical analysis was the chosen method to examine the relationship between inputs to and outputs of schooling in this study.

Method and Results

The purpose of the study was two-fold. First, researchers sought to confirm the results from two analytic techniques, namely regression and canonical correlation. Second, by using a method of analysis that would accommodate multiple output measures, researchers sought to more fully explain the relationship between inputs to schooling and measures of student achievement. Toward that end, a comparison of results from multiple regression and canonical analysis are presented.

Table 1
Descriptive Statistics

Variable	Mean	Standard Deviation
LEP	.69	.933
FREERED	39.261	17.0040
SPED	11.257	3.5558
MAJMIN	98.07	3.79
PCTPD	98.39	8.305
MASTERS	76.93	9.419
AVE_YEARS_EXP	11.902	1.842
SPENDING	5,310.45	1,210.969
STRATIO	17.04	2.084
ST_COMP_RATIO	4.405	1.5209
CTBSLANG	50.06	10.618
KCCTWR	64.54	8.827
RETAINED	6.008	3.4737
DROUPOUT	2.973	1.693
COLLEGE	52.854	14.8040
MILITARY	2.873	1.8345
WORKFORCE	27.801	11.5170
VOCED	5.302	3.9005
PARTTIME	6.882	7.1589
FAILURE	3.769	2.9757
n = 193		

Sampling and Participants

The choice of both independent and dependent variables was guided by a review of current literature. The study made use of school level data from the 2003–04 academic year. Data were collected from 193 high schools serving students in grades 9 through 12 across the Commonwealth of Kentucky. Descriptive statistics are displayed in Table 1.

Independent Variables

Stiefel, Schwartz, Rubenstein, and Zabel state that studies attempting to discern the relationship between resources and student achievement have included student demographics, resources, and organizational characteristics as independent variables.³⁶ By controlling for variables out of the control of the educational institution, such as student characteristics, efficiency measurements provide an opportunity to identify successful schools – especially schools where success may not be readily apparent. Measures of student attributes included in this study were, the percentage of students who received free and reduced lunch, the percentage of students who received special education services, and the percentage of students who received limited English proficiency services.

Current research has clearly identified the teacher as the single most important school-related input to improve student achievement.³⁷

Researchers, economists, and policy makers have made use of education production functions in an attempt to determine the relationship between teacher quality and student achievement.³⁸ These studies employed measurable, policy-relevant variables to describe teacher quality such as teacher certification, performance on certification exams, years of experience, relationship of teaching assignment to college major, teacher education level and student-teacher ratio.³⁹ Accordingly, this study included multiple measures of teacher quality as inputs to schooling. Included in the list were percent of teachers with a major or minor in the content area taught, percent of teachers participating in professional development, education level of the teacher as measured by the percentage of teachers holding a masters degree, and average years of experience.

The input variable per pupil expenditure was included in this study. This variable is often included in input-output studies although findings are mixed.⁴⁰ The negative relationship found to exist between per pupil expenditure and student achievement is likely the result of the additional cost of educating students in underrepresented populations or those with disabilities. While the literature clearly shows that all students can learn at high levels, the cost of providing needed services may be influenced by student need, concentration of need, and school location.⁴¹ Class size is an input variable that has been found to impact student achievement.⁴² That variable was included in this study and was defined as the average number of students in each class in the school for each teacher.

Student-computer ratio was a final variable included in the study. Jones and Paolucci argue that the exponential increase in expenditures on technology in K-12 schools and institutions of higher education make this variable increasing important to researchers.⁴³ Further, the acquisition of skills in the use of technology is an area of focus of standards based reform as states have begun to incorporate technology in to the curriculum so that student transition from school to work may be enhanced.⁴⁴ Using data from NAEP testing, Wenglinsky examined the relationship between computer use and student achievement.⁴⁵ He found that the largest impact on student achievement was made by teachers who used technology to promote higher order thinking skills. Further, his study suggested that time spent working on school related work at home was related to student achievement thus raising the question of access to and availability of technology. This issue is important in Kentucky given the prevalence of poverty in the state and given the fact that students experiencing poverty have been shown to lag behind their more affluent peers in computer use.⁴⁶

Dependent Variables

The 2004 Commonwealth Accountability Testing System (CATS) index was the dependent variable used in the multiple regression analysis. CATS recognizes the myriad purposes of education and makes use of multiple measures of student performance including the criterion referenced Kentucky Core Content Test (KCCT), a nationally norm-referenced test (e.g., the CTBS/5 Survey Edition), writing portfolios, and non-academic performance data (e.g., attendance, retention, and dropout rates; student transitions to next level of schooling and to adult life). Performance on each of these measures is differentially weighted to calculate a Kentucky Accountability Index for each school. Proficiency has been defined as an index score of 100. All schools are required to reach proficiency by 2014. CATS index scores are calculated yearly, although the system of sanctions and recognition operates on a biennial calendar.

To make the comparison between the multiple regression analysis and the canonical analysis unbiased, the components of the 2004 CATS index were used as the multiple dependent variables in the canonical analysis. Due to problems of multicollinearity, not all norm-referenced and criterion-referenced measures of student achievement could be used in the analysis. Researchers selected the norm-referenced test that had the smallest Pearson correlation with one of the criterion-referenced tests. This decision was made to preserve the integrity of the model because multicollinearity causes an inflated relationship in canonical analysis. The CTBS reading test was chosen as the norm-referenced test while the KCCT writing index was chosen as the criterion-referenced measure for inclusion in the canonical analysis. All non-academic measures of student achievement that comprise the CATS index were included in the canonical analysis. These measures included: percent of students retained, percent of students who were classified as dropouts, percent of students transitioning to college, percent of students entering the military, percent of students entering the workforce from high school, percent of students enrolling in vocational education, percent of students attending school part-time and working part-time, and percent of students who failed to make a successful transition following high school. Descriptive statistics appear in Table 1.

Guidelines for Interpretation

Sheskin⁴⁷ and Thompson⁴⁸ state the complexity of calculation coupled with the difficulty of interpretation of results has limited the use of canonical analysis. As such, a brief explanation of guidelines for interpretation is offered. First, the statistical significance of each canonical correlation is determined by a Wilk's test of significance. Interpretation of these results is similar to that of a Pearson correlation as one is interested in significance, size, and total variance explained by each relationship. The researcher retains any canonical correlations that are found to be statistically significant and proceeds to interpret any statistics (canonical loadings, standardized canonical coefficients, and cross loadings) that are associated with the canonical variates. Finally, the examination may include an inspection of redundancy. Unlike multiple regression which limits the interpretation of prediction to the relative importance of independent variables, three types of analysis are possible using canonical analysis. These include an interpretation of the relative importance of independent variables, an interpretation of the relative importance of dependent variables, and an interpretation of the relationship of individual variables with the linear combination of variables in the opposite set.

Both the standardized canonical coefficients and the canonical loadings provide the necessary information to discern the relative importance of independent and dependent variables. Standardized canonical coefficients are weights assigned to each variable so that the maximum possible Pearson correlation can be found between the canonical variates. The use of the standardized canonical coefficients is valuable since the coefficients are partial coefficients with the effect of the other variables removed.⁴⁹ Standardized canonical coefficients are interpreted in much the same way that one interprets a standardized regression coefficient in multiple regression.

The correlation between the canonical variate and the variable is called the canonical loading. The cross loading is the correlation between individual variables and the linear combination of the opposite set of variables. During each of these examinations, the researcher is interested in the largest (absolute value) coefficients or correlations that

are used.⁵⁰ The literature reveals that an interpretation of the results of canonical analysis is strengthened by an examination of canonical loadings and cross loadings for two reasons. First, it is assumed that there is greater stability in the correlation statistic when there are high or fairly high intercorrelations among the variables and the sample is of small or medium size. Second, the correlations provide a clearer indication of which variables are most closely aligned with the canonical variate. The researcher is interested in these correlations since the canonical variate is an unobserved trait.⁵¹ As a rule of thumb, canonical loadings and cross loadings that are greater than .30 should be treated as meaningful.⁵²

Redundancy in canonical analysis is the proportion of the variance in the X's that are predicted from, or explained by the linear combination of Y's. Redundancy is typically only calculated for canonical variates from statistically significant canonical correlations and these calculations are made based on the research design.⁵³ When predictor and criterion variables are used, the redundancy calculation is only made for the criterion variables since one is interested in determining the

proportion of the variance that is predictable. It is important to note that redundancy is not a measure of multivariate association and that this calculation will differ from the total amount of variance explained by the linear combination of variables.

Results of the Sequential Multiple Regression

A sequential multiple regression was performed using the 2004 CATS index as the dependent variable. Independent variables were entered in two blocks. The first block included student demographic data. Input variables in model 1 included the percent of students receiving services for limited English proficiency, the percent of students qualifying for free and reduced lunch, and the percent of students receiving services for special education. The second block of input variables included variables that were identified in the literature review that have been determined to have a relationship to student achievement. Those variables included percent of teachers holding a major or minor in the content area taught, percent of teachers who participated in professional development activities, percent of teachers holding an

Table 2
Multiple Regression Results

Table 2.1						
Model	Variables Entered	R	R Square	R Square Change	F Change	Significance of Change
1	LEP, FREERED, SPED	.779	.607	.607	97.386	.000
2	LEP, FREERED, SPED, ST_COMP_RATIO, PCTPD, MAJMIN, MASTERS, STRATIO, AVE_YEARS_EXP, SPENDING	.801	.642	.035	2.525	.017
Table 2.2						
		Unstandardized Coefficients		Standardized Coefficients		
Model	Variables Entered	B	Std Error	Beta	Significance	Tolerance
1	Constant	91.139	1.370			
	LEP	-.070	.203	-.016	.729	.990
	FREERED	-.300	.025	-.597	.000	.821
	SPED	-.748	.121	-.311	.000	.825
2	Constant	66.551	11.974			
	LEP	-.088	.203	-.020	.667	.931
	FREERED	-.295	.026	-.587	.000	.750
	SPED	-.733	.130	-.305	.000	.671
	MAJMIN	.236	.103	.105	.023	.948
	PCTPD	.039	.046	.038	.405	.968
	MASTERS	.104	.045	.114	.023	.787
	AVE_YEARS_EXP	.038	.241	.008	.874	.732
	SPENDING	-.001	.000	-.080	.156	.625
	STRATIO	-.464	.216	-.113	.033	.711
	ST_COMP_RATIO	-.058	.259	-.010	.823	.926

Table 3
Canonical Analysis Results with Demographic Student Data Input Only

Demographic Student Data Input	First Canonical Variate			Second Canonical Variate			Total
	Loading	Coefficient	Cross Loading	Loading	Coefficient	Cross Loading	
Inputs of Schooling:							
LEP	-.151	-.100	-.118	-.794	-.852	-.261	
FREERED	-.943	-.788	-.736	.248	.551	-.062	
SPED	-.679	-.356	-.530	-.325	-.576	.184	
Outputs of Schooling:							
CTBSREAD	.973	.968	.760	.072	.391	.024	
KCCTWR	.554	.075	.433	.040	.205	.013	
RETAINED	-.363	.210	-.283	-.347	-.428	-.114	
DROUPOUT	-.454	-.172	-.354	.126	.171	.042	
COLLEGE	.505	.072	.394	-.420	.350	-.138	
MILITARY	-.114	.043	-.089	.353	.294	.116	
WORKFORCE	-.489	.021	-.382	.421	.871	.139	
VOCED	-.261	-.023	-.203	-.210	.117	-.069	
PARTTIME	-.105	.027	-.082	.246	.423	.081	
FAILURE	-.312	.029	-.244	.568	.687	.187	
Canonical Correlation	.780			.329			
Wilk's (DF)	.321 (30)			.822 (18)			
Significance	.000			.007			
Percent of Variance	60.8			10.8			71.6
Redundancy	13.9			1.1			15.0

advanced degree (masters), average years of teaching experience, spending per pupil, student-teacher ratio, and student-computer ratio. Sequential multiple regression was the chosen method of analysis so that variance explained by student demographic could be separated from the variance explained by inputs to schooling so that efficiency conclusions could be drawn.

Results from the sequential multiple regression are presented in Table 2. According to those data, student demographics significantly predict student achievement in model 1, $R^2=.607$, $R^2_{adj}=.601$, $F(3, 189)=97.386$, $p<.000$. Model 1 accounted for 60.7% of the variance in student achievement as measured by the 2004 CATS index. Table 2 also displays the unstandardized regression coefficients (B), standardized regression coefficients (β), significance level of the regression coefficients, and tolerance for each independent variable. These data

enable the researcher to discern which independent variables were significant predictors of student achievement. Individually, the independent variables percent of students receiving special education services ($t=-6.193$, $p<.000$) and percent of students receiving free and reduced lunch ($t=-11.859$, $p<.000$) significantly predicted student achievement in model 1 as measured by the 2004 CATS index. Measures of tolerance calculated in the model indicated that multicollinearity was not a problem. Model 2 in the sequential multiple regression was also found to be a significant predictor of student achievement, $R^2=.642$, $R^2_{adj}=.622$, $F(7, 182)=2.525$, $p<.017$. Model 2 accounted for an additional 3.5% of the variance. Total variance explained in the regression analysis was 64.2% of the variance in student achievement. Input variables that were found to be significant predictors of student achievement in model 2 included percent of students receiving special education services

($t=-5.628$, $p<.000$), percent of students receiving free and reduced lunch ($t=-11.466$, $p<.000$), percent of teachers with a major or minor in the content area ($t=2.295$, $p<.023$), percent of teachers with an advanced degree (masters) ($t=2.287$, $p<.023$), and student-teacher ratio ($t=-2.148$, $p<.033$). Measures of tolerance revealed that multicollinearity was not a problem in the model.

Results of the Canonical Analysis

Unlike multiple regression, canonical analysis does not allow the researcher to control for covariance. In order to compare the results of the multiple regression analysis with the results from canonical analysis, two separate canonical analyses were calculated. Similar to model 1 in the multiple regression analysis, the only input variables included in the first canonical analysis were student demographics. The second canonical analysis included all input variables to detect any changes in the explained variance for the dependent variables. Results from the second canonical analysis were compared with model 2 in the multiple regression.

Results from the first canonical analysis are displayed in Table 3. Wilk's test of significance revealed that two canonical correlations computed in the first canonical analysis were significant ($R_c=.780$, Wilk's (30)=.321, $p<.000$; $R_c=.329$, Wilk's (18)=.822, $p<.007$, respectively). The first variate pair accounted for 60.8% of the total variance. The second variate pair accounted for 10.8% of the variance. Total pooled variance for this model is 71.6%. Using the aforementioned guidelines for interpretation, researchers identified independent variables that were deemed to be of importance, dependent variables that were deemed to be of importance, and interpreted the relationship between individual variables and the linear combination of the opposite set of variables. Independent variables that were deemed important in the first canonical variate included: the percentage of students receiving services for free and reduced lunch (*canonical coefficient*=-.788) and percentage of students receiving services for special education (*canonical coefficient*=-.356). Dependent variables that were deemed important in the first canonical variate included scores on the CTBS reading test (*canonical coefficient*=.968). An important relationship was found to exist between the independent variables percentage of students receiving services for free and reduced lunch (*canonical loading*=.736) and percentage of students receiving services for special education (*canonical loading*=.530) and the linear combination of dependent variables in the first canonical variate. Finally, an important relationship was found to exist between the dependent variables scores on the CTBS reading test (*canonical loading*=.760), scores on the KCCT writing test (*canonical loading*=.433), percentage of dropouts (*canonical loading*=-.354), percentage of students enrolling in a four year college (*canonical loading*=.394), and percentage of students entering the workforce (*canonical loading*=.382).

Results from the second canonical variate identified a third measure of student demographics as an important predictor of student achievement. In addition to the percentage of students receiving services for free and reduced lunch (*canonical coefficient*=-.852) and percentage of students receiving services for special education (*canonical coefficient*=-.551), the percentage of students receiving services for limited English proficiency (*canonical coefficient*=-.576) was found to be of relative importance to the relationship between student demographics and measures of student achievement. Further, the second canonical variate identified additional dependent measures of importance. In addition to scores on the CTBS reading test (*canonical coefficient*=.391),

percentage of students retained (*canonical coefficient*=-.428), percentage of students enrolling in a four year college or university (*canonical coefficient*=.350), percentage of students entering the workforce (*canonical coefficient*=.871), and percentage of students classified as working part time and attending school part time (*canonical coefficient*=.423) were identified as relatively important outputs of schooling. None of the cross loadings met the criteria of <.30 in the second canonical variate. As such, no additional important relationships were identified.

Results from the second canonical analysis are presented in Table 4. Wilk's test of significance revealed that two canonical correlations computed in the second canonical analysis were significant ($R_c=.799$, Wilk's (100)=.321, $p<.000$; $R_c=.435$, Wilk's (81)=.822, $p<.017$, respectively). The first variate pair accounted for 63.8% of the total variance. The second variate pair accounted for 18.9% of the variance. Total pooled variance for this model is 82.7%. Using the guidelines for interpretation, researchers identified independent variables that were deemed to be of importance, dependent variables that were deemed to be of importance, and interpreted the relationship between individual variables and the linear combination of the opposite set of variables. Independent variables that were deemed important in the first canonical variate included: the percentage of students receiving services for free and reduced lunch (*canonical coefficient*=.729) and percentage of students receiving services for special education (*canonical coefficient*=.352). Dependent variables that were deemed important in the first canonical variate included scores on the CTBS reading test (*canonical coefficient*=-.982). An important relationship was found to exist between the independent variables percentage of students receiving services for free and reduced lunch (*cross loading*=.703), percentage of students receiving services for special education (*cross loading*=.535), and spending per pupil (*cross loading*=.425) and the linear combination of dependent variables in the first canonical variate. Finally, an important relationship was found to exist between the dependent variables scores on the CTBS reading test (*cross loading*=-.786), scores on the KCCT writing test (*cross loading*=-.452), percentage of students retained (*cross loading*=.313), percentage of dropouts (*cross loading*=.332), percentage of students enrolling in a four year college (*cross loading*=-.385), and percentage of students entering the workforce (*cross loading*=.371).

Results from the second canonical variate identified four important input variables: percentage of students receiving services for limited English proficiency (*canonical coefficient*=-.650), percentage of teachers participating in content-focused professional development (*canonical coefficient*=.415), spending per pupil (*canonical coefficient*=-.479) and student teacher ratio (*canonical coefficient*=-.440). Further, the second canonical variate identified additional dependent measures of importance. In addition to scores on the CTBS reading test (*canonical coefficient*=.797), and percentage of students enrolling in a vocational school (*canonical coefficient*=.359) were identified as relatively important outputs of schooling. None of the cross loadings met the criteria of <.30 in the second canonical variate. As such, no additional important relationships were identified.

Discussion

The purpose of this study was to compare multiple regression with canonical analysis in order to introduce a new, policy relevant methodology to the literature on production functions. Findings from this study confirmed the results of past inquiries that found a relationship

Table 4
Canonical Analysis Results with All Input Variables

All Input Variables	First Canonical Variate			Second Canonical Variate			Total
	Loading	Coefficient	Cross Loading	Loading	Coefficient	Cross Loading	
Inputs of Schooling:							
LEP	.161	.109	.129	-.647	-.650	-.281	
FREERED	.913	.729	.703	.157	.272	.068	
SPED	.669	.352	.535	-.087	-.077	.038	
MAJMIN	-.253	-.092	-.202	-.181	-.186	-.079	
PCTPD	-.085	-.082	-.068	.337	.415	.146	
MASTERS	-.049	-.016	-.039	-.421	-.264	-.183	
AVE_YEARS_EXP	-.332	-.078	-.265	-.107	-.005	-.046	
SPENDING	.532	.140	.425	-.332	-.479	-.145	
STRATIO	-.304	.171	-.243	-.216	-.440	-.094	
ST_COMP_RATIO	-.071	-.036	-.056	-.125	-.039	-.054	
Outputs of Schooling:							
CTBSREAD	-.983	-.982	-.786	.068	.748	.030	
KCCTWR	-.566	-.100	-.452	-.196	-.059	-.085	
RETAINED	.392	-.153	.313	-.239	-.240	-.104	
DROUPOUT	.415	.103	.332	.210	.295	.091	
COLLEGE	-.482	-.058	-.385	-.668	-.169	-.291	
MILITARY	.108	-.040	.086	.243	.193	.106	
WORKFORCE	.465	-.050	.371	.628	.797	.273	
VOCED	.267	.023	.214	.149	.359	.065	
PARTTIME	.102	-.024	.081	.178	.180	.078	
FAILURE	.303	-.030	.242	.254	.284	.110	
Canonical Correlation	.799			.435			
Wilk's (DF)	.197 (100)			.544 (81)			
Significance	.000			.017			
Percent of Variance	63.8			18.9			82.7
Redundancy	14.4			2.2			16.6

between the inputs to schooling and measures of student achievement. A statistically significant relationship was found to exist through the use of canonical analysis. For the purpose of this discussion, we focus on the findings from the second canonical analysis. That model made use of ten independent variables and ten dependent measures of student achievement. Two of the ten canonical correlations calculated revealed a statistically significant relationship. Together, the pooled variance explained 82.7% of the variance between inputs to schooling and measures of student achievement. By using multiple measures of student achievement, the chosen method of analysis enabled researchers to explain a greater percentage of variance than was explained through the use of multiple regression. As suggested in the literature review, schools produce multiple outcomes; therefore the selection of a method of analysis that allowed for the interaction of all of those variables in a linear combination of output variables allowed researchers to more fully explain the relationship between inputs to schooling and measures of student achievement.

The use of canonical analysis confirmed that student demographics, as identified in the multiple regression, are significant predictors of student achievement. Because interpretations of canonical loadings, standardized canonical coefficients, and cross loadings make use of absolute values conclusions with regard to the direction of the relationship are not possible. The method of analysis enabled the identification of all three measures of student demographics as important. Through the use of multiple regression, limited English proficiency (LEP) was not identified as a significant predictor of student achievement even though policy implications about LEP abound. Given the small percentage of students identified as limited English proficiency in the Commonwealth of Kentucky, the finding of a relationship is significant and has policy implications. The use of canonical analysis has allowed for the interaction of multiple outputs of schooling and therefore aided in the identification of an area for further research and intervention.

Aside from measures of student demographics, multiple input resources were found to be significant predictors of student achievement through the use of canonical analysis. The multiple regression analysis identified the variables major or minor in the content area, education level of teachers (master's degree) and student teacher ratio as significant predictors of student achievement. By using canonical analysis, researchers found that spending per pupil, student-teacher ratio, and percent of teachers participating in content focused professional development were significant predictors of student achievement. Professional development is not a variable that has been found to be a significant predictor of student achievement in the literature. This study has identified that variable as an area of future inquiry. Most importantly, this study clearly links the input resources with measures of student achievement making this method of analysis a viable method for the study of resource efficiency.

The main difference between multiple regression and canonical analysis is that the researcher may make use of multiple dependent measures. Because schools produce multiple outputs, it has been postulated that this method of analysis better enables the researcher to simulate reality. The use of multiple output measures eliminates researcher bias. This methodology does not require the researcher to choose one independent measure. Results from this study indicated that the most important output of schooling, given the ten dependent measures, was reading. The identification of literacy as the predominant output of schools has tremendous policy implications

when one considers state and national goals with regard to access to and completion rates of higher education to drive the economy. Further, the identification of workforce entry and percentage of students enrolling in vocational schools as important outputs of schooling is noteworthy in a time of standards based reform. Without casting dispersions on the current movement of educational reform, it is undeniable that the focus on standards and student achievement as measured by standardized testing may have disillusioned students from pursuing these interests. The production of academic skills has been the priority of public schools of late. As such, schools have had to cut back on programs such as vocational education and tech prep. These findings suggest that schools produce more than just academic results and that a focus on vocational programs has merit in our high schools so long as the proper counseling is provided to students with regard to life opportunity and so that students are not categorized and tracked based on ethnicity or socioeconomic status. All children must be afforded the equal opportunity to pursue their own educational and occupational goals.

Results from this study are important for both policymakers and practitioners because they suggest the need for an alignment of educational practice. Schools make use of a variety of resources to achieve multiple goals. The realization of these sometimes competing goals requires an educational leader with the vision, knowledge dispositions, and leadership skills to align the school mission with research based educational best practice in order to maximize student achievement, however that is defined. Schools cannot afford to focus their energies on one specific goal or one subpopulation in the entire student body. Current educational policy that requires proficiency for all coupled with the realities of globalization and increased international competition necessitate a rethinking of the focus and leadership of schools. Empirical research must include these multiple contingencies to help inform practice. Canonical analysis is one method with the potential to do that.

A limitation of this study was that data were aggregated to the school level and included merely one year's worth of data. While acknowledging the limitations of this data set, this study has identified canonical analysis as a methodology that more fully explains the relationship between input resources to schooling and multiple output measures. We envision an extension of this study wherein a canonical correlation is calculated for each individual school. The myriad of ways by which results from canonical analysis may be interpreted enable the researcher to examine not just important inputs to schooling but also to identify the outputs of importance at each school and the interaction of all variables. The ability to examine the outputs of schools has merit given current educational policy. With proficiency goals looming by 2014 for both state and national education policy, canonical analysis may identify the need to change both focus and practice at the school level so that policy goals of social justice may be obtained. We envision these results being useful by policymakers and educational leaders who must confront the belief systems of practitioners with regard to what and how much students from different socioeconomic and ethnic groups can learn.

The redundancy statistic is included in the analysis to temper the size of the relationship that was found in this study. The research clearly states that the redundancy statistic is not to be used as an analytical technique. For the purposes of this study, the redundancy statistic demonstrates that the predictive model presented in this study can be used to discern the relationship between inputs to schooling

and measures of student achievement. Total redundancy in the model was 16.6% which suggests that the inputs utilized in this study are predictors of student achievement. Moreover, it suggests that the model has not accounted for all factors that are present in the relationship between inputs to schooling and measures of student achievement. In examining the relationship between measures of teacher quality and student achievement, Rice notes that the research has been limited to policy relevant, measurable variables.⁵⁴ Results from this study suggest the need for more and better variables at the classroom level that more fully capture the process of teaching and learning. Not only do we as researchers need better sets of data that disaggregate data at the classroom level, we need to develop better tools to measure student-teacher interaction, communication, teacher reflection, and the use of assessment measures in the educational process. By more fully capturing the ability to measure the educational process, research becomes more relevant for educational leaders who seek to maximize student achievement.

Endnotes

¹ This article is based upon a paper originally presented at the Annual Conference of the American Education Finance Association, March 2007, Baltimore, Maryland.

² Leanna Stiefel, Amy E. Schwartz, Ross Rubenstein, and Jeffrey Zabel, "Measuring School Efficiency: What Have We Learned?" in *Measuring School Performance and Efficiency: Implications for Practice and Research*, ed. Leanna Stiefel, Amy E. Schwartz, Ross Rubenstein, and Jeffrey Zabel, 2005 yearbook of the American Education Finance Association (Larchmont, NY: Eye on Education, 2005), 1-15.

³ Michael A. Rebell and Joseph J. Wardeniski, *Of Course Money Matters: Why the Arguments to the Contrary Never Added Up* (New York: Campaign for Fiscal Equity, 2004); Deborah A. Verstegen and Richard A. King, "The Relationship Between School Spending and Student Achievement: A Review and Analysis of 35 Years of Production Function Research," *Journal of Education Finance* 24 (Fall 1998): 241-262; Helen F. Ladd and Janet S. Hansen, ed., *Making Money Matter: Financing America's Schools* (Washington, D.C.: National Academy Press, 1999).

⁴ David H. Monk, "The Education Production Function: Its Evolving Role in Policy Analysis," *Educational Evaluation and Policy Analysis* 11 (Spring 1989): 31-45.

⁵ Amy E. Schwartz, Leanna Stiefel, and Hella B. Hadj Amor, "Measuring School Performance Using Cost Functions," in *Measuring School Performance and Efficiency*, 67-92.

⁶ Ross Rubenstein, "The Reliability of School Efficiency Measures Using Data Envelopment Analysis," in *Measuring School Performance and Efficiency*, 93-118.

⁷ Robert C. Knoepfel, Deborah A. Verstegen, and James S. Rinehart, "What Is the Relationship Between Spending and Student Achievement? A Canonical Analysis," *Journal of Education Finance* 33 (Fall 2007): 183-202.

⁸ Anthony Rolle, "Out with the Old – In with the New: Thoughts on the Future of Educational Productivity Research," *Peabody Journal of Education* (July 2004): 31-56.

⁹ Stiefel et al., "Measuring School Efficiency," in *Measuring School*

Performance and Efficiency; Amy E. Schwartz and Jeffrey Zabel, "The Good, the Bad, and the Ugly: Measuring School Efficiency Using School Production Functions," in *Measuring School Performance and Efficiency*, 37-66; Monk, "The Education Production Function."

¹⁰ Schwartz and Zabel, "The Good, the Bad, and the Ugly."

¹¹ Barbara G. Tabachnick and Linda S. Fidell, *Using Multivariate Statistics*, 5th ed. (Boston: Pearson Education, Inc., 2007); Elazar J. Pedhazur, *Multiple Regression in Behavioral Research: Explanation and Prediction*, 3d ed. (New York: Harcourt Brace College Publishers, 1997).

¹² Tabachnick and Fidell, *Using Multivariate Statistics*; Pedhazur, *Multiple Regression in Behavioral Research*, 3d ed.; James P. Stevens, *Applied Multivariate Statistics for the Social Sciences*, 3d ed. (Mahwah, NJ: Lawrence Erlbaum Associates, Publishers, 1996).

¹³ Pedhazur, *Multiple Regression in Behavioral Research*, 3d ed.

¹⁴ Jacob Cohen and Patricia Cohen, *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 2d ed. (Hillsdale, NJ: Lawrence Erlbaum Associates, 1983).

¹⁵ Cohen and Cohen, *Applied Multiple Regression/Correlation Analysis*; Pedhazur, *Multiple Regression in Behavioral Research*, 3d ed.; Stevens, *Applied Multivariate Statistics*; Tabachnick and Fidell, *Using Multivariate Statistics*.

¹⁶ James Coleman, Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York, *Equality of Educational Opportunity*, Vol. 1 (Washington, DC: U.S. Department of Health, Education, and Welfare, 1966).

¹⁷ Deborah A. Verstegen, "Efficiency and Equity in the Provision and Reform of American Schooling," *Journal of Education Finance* 20 (Summer 1994): 107-131.

¹⁸ Jim C. Fortune and John S. O'Neil, "Production Function Analysis and the Study of Educational Funding Equity: A Methodological Critique," *Journal of Education Finance* 20 (Summer 1994): 21-46; Deborah A. Verstegen, "The American Education Reform Movement," *Journal of Educational Planning and Administration* 8 (July 1994): 281-295.

¹⁹ Fortune and O'Neil, "Production Function Analysis."

²⁰ Eric A. Hanushek, "The Impact of Differential Expenditures on School Performance," *Educational Researcher* 18 (May 1989): 45-51; Eric A. Hanushek, "When School Finance Reform May Not Be Good Policy," *Harvard Journal of Legislation* 28 (Summer 1991): 465-498.

²¹ Larry V. Hedges, Richard D. Laine, and Rob Greenwald, "When Reinventing the Wheel Is Not Necessary: A Case Study in the Use of Meta-Analysis in Education Finance," *Journal of Education Finance* 20 (Summer 1994): 1-20.

²² Hanushek, "The Impact of Differential Expenditures"; Hanushek, "When School Finance Reform May Not Be Good Policy."

²³ Hedges et al., "When Reinventing the Wheel Is Not Necessary."

²⁴ Monk, "The Education Production Function."

²⁵ Ibid.

²⁶ Pedhazur, *Multiple Regression in Behavioral Research*, 3d ed.; Tabachnick and Fidell, *Using Multivariate Statistics*; Faith E. Crampton and Terry N Whitney, *The Relationship Between Educational Expenditure and Student Achievement: When Does Money Matter?* Working Paper (Denver, CO, and Washington, DC: National Conference of

State Legislatures, 1996), 1-11; Cohen and Cohen, *Applied Multiple Regression*; Elazar J. Pedhazur, *Multiple Regression in Behavioral Research: Explanation and Prediction*, 2d ed. (New York: Holt, Rinehart, and Winston, 1982).

²⁷ Bruce Thompson, "Methods, Plainly Speaking: A Primer on the Logic and Use of Canonical Correlation Analysis," *Measurement and Evaluation in Counseling and Development* 24 (July 1991): 80-93.

²⁸ Ibid.

²⁹ Pedhazur, *Multiple Regression in Behavioral Research*, 3d ed.; Thompson, "Methods, Plainly Speaking."

³⁰ Thompson, "Methods, Plainly Speaking."

³¹ Pedhazur, *Multiple Regression in Behavioral Research*, 3d ed.; Pedhazur, *Multiple Regression in Behavioral Research*, 2d ed.

³² Francis J. Kelly, Donald L. Beggs, and Keith A. McNeil, *Research Design in the Behavioral Sciences: Multiple Regression Analysis* (Carbondale, IL: Southern Illinois University Press, 1969).

³³ Pedhazur, *Multiple Regression in Behavioral Research*, 3d ed.; Pedhazur, *Multiple Regression in Behavioral Research*, 2d ed.

³⁴ Monk, "The Education Production Function."

³⁵ Sam Kachigan (1986) as cited in D.J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, 2d ed. (Washington, DC: Chapman & Hall/CRC, 2001).

³⁶ Stiefel et al., "Measuring School Efficiency," in *Measuring School Performance and Efficiency*.

³⁷ Gordon Cawelti, "Improving Achievement: Finding Research-Based Practices and Programs that Boost Student Achievement," *The American School Board Journal* 186 (July 1999): 34-37; H. Jordan, Robert Mendro, and Dash Weerasinghe, "Teacher Effects on Longitudinal Student Achievement," paper presented at the Sixth Annual Evaluation Institute sponsored by CREATE, Indianapolis, July 1997; Leslie S. Kaplan and William A. Owings, "The Politics of Teacher Quality: Implications for Principals," *NASSP Bulletin* 86 (December 2002): 22-41.; William L. Sanders and June C. Rivers, "Cumulative and Residual Effects of Teachers on Future Student Academic Achievement," Research Progress Report (Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center, 1996); S. Paul Wright, Sandra P. Horn, and William L. Sanders, "Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation," *Journal of Personnel Evaluation in Education* 11 (April 1997): 57-67; Linda Darling-Hammond, "Quality Teaching: The Critical Key to Learning," *Principal* 77 (September 1997): 5-6.

³⁸ Jennifer King Rice, Teacher Quality: Understanding the Effectiveness of Teacher Attributes, (Washington, DC: Economic Policy Institute, 2003).

³⁹ Steven G. Rivkin, Eric A. Hanushek, and John F. Kain, "Teachers, Schools and Academic Achievement," *Econometrica* 73 (March 2005): 417-458; Rebell and Wardenski, "Of Course Money Matters."; Alan B. Krueger, "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics* 114 (May 1999): 397-532; Deborah A. Verstegen, Richard A. King and Bettye MacPhail-Wilcox, "What We Know About the Effect of Resources on Student Cognitive Achievement from Thirty Years of Production Function Analysis," paper presented at the Annual Conference of the American Education

Finance Association, Jacksonville, FL, March 1997; Hedges et al., "When Reinventing the Wheel Is Not Necessary"; David H. Monk, "Subject Area Preparation of Secondary Mathematics and Science Teachers and Student Achievement," *Economics of Education Review* 13 (June 1994): 31-45; Hanushek, "The Impact of Differential Expenditures"; Hanushek, "When School Finance Reform May Not Be Good Policy."

⁴⁰ Hedges et al., "When Reinventing the Wheel Is Not Necessary"; Verstegen et al., "What We Know About the Effect of Resources on Student Cognitive Achievement"; Hanushek, "The Impact of Differential Expenditures"; Hanushek, "When School Finance Reform May Not Be Good Policy."

⁴¹ Bruce D. Baker, "The Emerging Shape of Educational Adequacy: From Theoretical Assumptions to Empirical Evidence," *Journal of Education Finance* 30 (Winter 2006): 259-287.; Richard Rothstein, *Class and Schools* (Washington, DC: Economic Policy Institute, 2004); David C. Berliner, "Our Impoverished View of Educational Reform," *Teachers College Record* 108 (June 2006): 949-995, <http://www.tcrecord.org/content.asp?contentid=12106>; Valerie E. Lee and David T. Burkham, *Inequality at the Starting Gate: Social Background Differences in Achievement as Children Begin School* (Washington, DC: Economic Policy Institute, 2002); Deborah A. Verstegen, "The New Finance," *The American School Board Journal* 189 (October 2002): 24-26.

⁴² Rivkin et al., "Teachers, Schools and Academic Achievement."; Rebell and Wardenski, "Of Course Money Matters"; Barbara Nye, Larry V. Hedges, and Spyros Konstantopoulos, "The Long-Term Effects of Small Classes: A Five-Year Follow-Up of the Tennessee Class Size Experiment," *Education Evaluation and Policy Analysis*, 21 (June 2000): 127-142.; Krueger, "Experimental Estimates of Education Production Functions"; Verstegen et al., "What We Know About the Effects of Resources"; Jeremy D. Finn and Charles M. Achilles, "Answers and Questions About Class Size: A Statewide Experiment," *American Educational Research Journal* 27 (September 1990): 557-577; Word et al., *Student Teacher Achievement Ratio (STAR)*.

⁴³ Trevor H. Jones and Rocco Paolucci, "Research Framework and Dimensions for Evaluating the Effectiveness of Educational Technology Systems on Learning Outcomes," *Journal of Research on Computing in Education* 32 (Fall 1999): 17-27.

⁴⁴ Harold Wenglinsky, *Using Technology Wisely: The Keys to Success in Schools* (New York: Teachers College Press, 2005).

⁴⁵ Harold Wenglinsky, "Technology and Achievement: The Bottom Line," *Educational Leadership* 63 (December 2005/January 2006): 29-32.

⁴⁶ Henry J. Becker, "Who's Wired and Who's Not: Children's Access to and Use of Computer Technology," *Children and Computer Technology* 10 (Fall/Winter 2000): 44-75.

⁴⁷ David J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, 2d ed., (Washington, DC: Chapman & Hall/CRC, 2001).

⁴⁸ Thompson, "Methods, Plainly Speaking."

⁴⁹ Stevens, *Applied Multivariate Statistics*.

⁵⁰ Ibid.

⁵¹ Weinberg and Walberg, (1973), as cited in Stevens, *Applied Multivariate Statistics*.

⁵² Pedhazur, *Multiple Regression in Behavioral Research*, 2d ed.

⁵³ Ibid.

⁵⁴ Rice, "Teacher Quality."

Funding Michigan K-12 Adequacy Without Rewarding Inefficiency

James J. Walters

Taxpayers and politicians expect public schools to exercise stewardship and wisdom regarding the use of resources entrusted to them. These public expectations approximate what economists refer to as technical efficiency. Technical efficiency emerges from the ideal use of available resources for maximizing output whereas allocative efficiency derives from comparing alternative technically efficient systems and choosing the least costly option.¹ A third and more obscure type of efficiency emerges in economic analysis from an interpretation of the unobserved effects of the entity studied. This phenomenon is sometimes referred to as "x-efficiency." Its significance comes from the unobserved effects of vision, motivation, incentives, and the culture of the entity and its leadership.²

Evidence exists that qualitative factors such as clearly defined goals, uninhibited access to information regarding these goals, incentives, motivation and effort, often the fruit of competition or adversity, yield far greater output improvement compared to marginal changes in inputs.³ Quantity times price may generate a variety of results depending on these unobserved factors. Improving student achievement by accomplishing changes in school organizational behavior represents direct application of x-efficiency.

The analysis in this study draws heavily on the notions of both technical efficiency and x-efficiency. Both of these lend themselves to an input/output style of inquiry like the education production function. This economic model builds on the foundation of the Cobb-Douglas factors of production theory although the genesis of that theory relates to industrial not educational formulations.⁴

Research Design

The goal of this study was to estimate the effects of district efficiency on student achievement in Michigan with the hope that objective analysis might serve to ease progress through the troublesome political process any transition to an adequacy-based school finance model will encounter. This study draws upon the methodology used by Phelps and Addonizio in their 2006 study of school accountability in Minnesota.⁵

Michigan does not track student achievement data by individual teacher or per pupil expenditures by school, only by district. Were per pupil expenditure available by school, the flow to individual students would require reliance on assumptions and averages. The unavailability of test score data by classroom or teacher, combined with the lack of reliable per pupil expenditure data by school and the abstraction caused by artificial resource flow assumptions, prompted the study's

use of the district as the unit of analysis. District level data for MEAP (Michigan Educational Assessment Program) scores and per pupil expenditure came from the State of Michigan website.

The operative version of the theoretical education production function for use in this study appears below:⁶

$$M_t = b_0 + b_1 ptenroll + b_2 avg_t_sal + b_3 avg_p_tchr + b_4 avg_isal + b_5 avg_totexp_ntr + u + e$$

Where

M represents statewide Michigan Education Assessment Program (MEAP) reading and math scores, stated as the percentage of students taking the test who achieved at a level meeting state standards;

ptenroll equals the percentage of students in a district eligible to receive free or reduced-price meals under U.S. federal guidelines;

avg_t_sal denotes the average teacher salary in the district;

avg_p_tchr is the average number of pupils per district teacher;

avg_isal is the average per pupil district expenditures related to instructional salaries;

avg_totexp_ntr controls for total district expenditures per pupil, net of transportation;

"*u*" signifies the portion of the residual that does not vary over time but does vary by district (This can be referred to as the district fixed effect and is estimated following regression);

"*e*" signifies the random portion of unobserved, residual, or unexplained variation.

Analysis of the residuals in the fashion indicated above requires retrieval of multiple observations for each district over time. This study includes a balanced panel of observations for districts over four years starting with the 2001-2002 school year through 2004-2005. The average residual by district was used to proxy for the district fixed effect in second stage regressions.

Although the model specified above contains no variable for district size, the regression technique used for this study was weighted by the full time equivalent student population for each district in each year. This adjusts for district size and mitigates the lack of constant variance in the residuals (heteroscedasticity) which represents one of the basic assumptions underlying linear regression.

Analysis of Data And Results

Data Description

Data were collected from public files available on the websites of the Michigan Department of Education (MDE) and Center for Educational Performance and Information (CEPI). Data for the dependent variable came from MEAP scores maintained by the Office of Educational Assessment and Accountability (OEAA) of the MDE. The second file type contained district financial information called Bulletin 1014 administered by the MDE Office of State Aid and School Finance. Data for student eligibility for federal meal subsidies came from information contained in the Single Record Student Data base controlled by CEPI. A file representing various measures of a single element in this database called Free and Reduced Lunch (FRL) appears on the CEPI website.

Bulletin 1014 files contained the most accurate district count as verified with the School Code Master file maintained by MDE. The

James J. Walters is a Certified Public Accountant and Chief Financial Officer for a group of private companies. He is also a part-time instructor in the Master's of Public Administration program at Western Michigan University.

number of districts reporting in Bulletin 1014 for the years included in the panel from 2001-2002 through 2004-2005 school years as follows: 743, 742, 744, and 760. However, only 494 districts reported data for every field used in the model for every year in the panel. The primary source for this discrepancy comes from counting each charter school as a separate district.⁷ However, several traditional districts were excluded from the study panel. Some traditional school districts in Michigan do not offer all twelve grades. For the study, any district that did not offer either seventh or eighth grade was necessarily eliminated from the panel. Also, MEAP scores are not reported in the public files for districts with fewer than ten test-takers in a grade.

Descriptive statistics for the 494 district panel are presented in Table 1. The summary of the dataset contained in Table 1 represents the same 494 Michigan school districts observed across four years for a total of 1,976 observations. The means and standard deviations reported for each explanatory variable were determined after weighting each variable by the inverse of variance for the student population. This technique is useful for observations containing averages. Averages based on the number of observations grow in precision as the number increases. Weighting provides the means to concede greater importance to the more precise measurements.⁸ Weighting considers the variation in the data by student although the unit of analysis remains aggregated by district.

Preliminary Annual Test Results

A preliminary set of sixteen regressions for all four measures of student achievement and separately for each of the four years served several purposes. Review of model specification, fit, and model diagnosis represented the primary motivation. The regressions were weighted by the student population of each school district as discussed above. This procedure corrected for the anticipated lack of constant variance in the model error term caused by the wide variance in district size as measured by the number of students. This heteroscedasticity

represented the principal diagnostic problem related to the underlying assumptions for least squares regression. The weighting methodology provided significant improvement but did not entirely correct the problem for all years in the study.⁹

Analysis of Residuals

Some variation in the student achievement measures from the regressions referred to above remained unexplained. These residuals contained the fixed but unobserved effect of the district plus random error.¹⁰ The average residual for each district was used to investigate systematic achievement above or below that predicted by the explanatory variables in each year. The result was assumed to measure the extent to which the district benefited from "x-efficiency," or contribution to student achievement not captured by the variables specified in the model. This estimate of district fixed effect was used as an explanatory variable in second stage regressions.

This simple averaging method for estimating district fixed effects was used after several attempts at fixed effects regression models failed to untangle the high correlation between the explanatory variables and fixed portion of the residual.¹¹ This correlation also proscribed the use of random effects or generalized least squares methodology.

Post Estimation Annual Test Results Including Fixed Effects Estimates

The sixteen regression results in Tables 2-5 came from estimating the same model described, but not presented, for preliminary annual tests, with one exception. The models estimated here included the variable determined in the previous section to represent the fixed effect of each district (avg_resid). This variable represented a relative measure of each district's contribution to the percentage of students meeting or exceeding state standards after controlling for the other predictors. The residual was averaged for each district using the results of the preliminary regressions for MEAP math and reading tests in fourth, seventh, and eighth grades. The results were analytically weighted by

Table 1
Descriptive Statistics, 2002-2005

Variables	Observations	Weight	Mean	Std. Dev.	Minimum	Maximum
district	1,976	6,438,484			1.010	83,070
year	1,976	6,438,484			2002	2005
math_gr4_sat	1,976	6,438,484	0.695465	0.143724	0.101	1
read_gr4_sat	1,976	6,438,484	0.741757	0.155611	0.13	1
read_gr7_sat	1,976	6,438,484	0.614158	0.174635	0.124	0.97
math_gr8_sat	1,976	6,438,484	0.572979	0.188827	0.057	1
pctenroll	1,976	6,438,484	0.333412	0.217264	0.02	0.9
avg_t_sal	1,976	6,438,484	54056.33	6903.321	24,547	83,479
avg_p_tchr	1,976	6,438,484	21.73831	2.565409	9	33
avg_isal	1,976	6,438,484	4663.104	585.9229	2,827	7,010
avg_totexp_ntr	1,976	6,438,484	8002.849	1294.894	5,416	15,628

Table 2
Grade 4 Math Scores Post-Estimation WLS Regression Results

Variables	School Year			
	2001–2002	2002–2003	2003–2004	2004–2005
pctenroll	-0.465***	-.0473***	-.0495***	-.0502***
	[-0.61]	[-0.63]	[-0.70]	[-0.75]
	(0.016)	(0.015)	(0.017)	(0.017)
avg_t_sal	-0.000000217	0.000000662	-0.000000596	0.000000939
	[-0.010]	[0.033]	[0-0.032]	[0.053]
	(0.00000063)	(0.00000059)	(0.00000060)	(0.00000058)
avg_p_tchr	-0.00331**	-0.00490***	-0.00720***	-0.00982***
	[-0.054]	[-0.088]	[-0.16]	[-0.22]
	(0.0014)	(0.0013)	(0.0014)	(0.0015)
avg_isal	0.0000564***	0.0000366***	0.0000473***	0.0000425***
	[0.20]	[0.13]	[0.19]	[0.19]
	(0.000011)	(0.0000095)	(0.0000095)	(0.000010)
avg_totexp_ntr	-0.0000216***	-0.0000177***	0.0000232***	-0.0000277***
	[-0.15]	[-0.13]	[-0.20]	[-0.26]
	(0.0000044)	(0.0000040)	(0.0000036)	(0.0000039)
avg_resid	1.036***	1.086***	1.125***	1.303***
	[0.60]	[0.63]	[0.69]	[0.83]
	(0.034)	(0.034)	(0.037)	(0.042)
Constant	0.907***	0.891***	1.010***	0.983***
	[6.92]	[6.81]	[8.22]	[8.30]
	(0.034)	(0.029)	(0.031)	(0.036)
Observations (n)	494	494	494	494
R-squared	0.90	0.90	0.89	0.87

Note: Normalized beta coefficients in brackets. Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

the inverse of variance for each district's student population. Each of the four tables of regression results presented represents one of the four measures of student achievement regressed over the independent variables for all four years included in the study.

The fixed effect variable (avg_resid) was statistically significant with a positive coefficient for all sixteen regressions. The measure for socioeconomic status (pctenroll) also remained statistically significant with a negative coefficient across all sixteen model iterations. A one percent increase in students eligible for free or reduced meals was

associated with anywhere from one-third to three quarters of a percent decrease in the percentage of students achieving state standards on the MEAP depending on the year and subject matter.

All the district resource variables except teacher salaries (avg_t_sal) were statistically significant for all of the regression models. The variable for teacher salaries remained statistically insignificant for all except two regressions. The pupil-teacher ratio (avg_p_tchr) was negative and statistically significant across all sixteen regressions. Its beta coefficient, with only one exception, represented the smallest impact

Table 3
Grade 4 Reading Scores Post-Estimation WLS Regression Results

	School Year			
Variables	2001–2002	2002–2003	2003–2004	2004–2005
pctenroll	-0.486***	-0.372***	-0.389***	-0.378***
	[-0.68]	[-0.63]	[-0.70]	[-0.75]
	(0.013)	(0.014)	(0.015)	(0.012)
avg_t_sal	-0.000000261	0.000000919*	-0.000000655	0.000000638
	[-0.013]	[-0.058]	[-0.045]	[-0.049]
	(0.00000051)	(0.00000054)	(0.00000052)	(0.00000041)
avg_p_tchr	-0.00593**	-0.00474***	-0.00527***	-0.00395***
	[-0.10]	[-0.11]	[-0.15]	[-0.12]
	(0.0011)	(0.0012)	(0.0012)	(0.0010)
avg_isal	0.0000475***	0.0000696***	0.0000419***	0.0000421***
	[0.18]	[0.33]	[0.22]	[0.25]
	(0.0000086)	(0.0000087)	(0.0000082)	(0.0000071)
avg_totexp_ntr	-0.0000302***	-0.0000376***	0.0000282***	-0.0000239***
	[-0.22]	[-0.35]	[-0.31]	[-0.30]
	(0.0000035)	(0.0000037)	(0.0000031)	(0.0000028)
avg_resid	0.978***	0.905***	0.799***	0.727***
	[0.60]	[0.67]	[0.63]	[0.62]
	(0.028)	(0.032)	(0.032)	(0.030)
Constant	0.991***	1.030***	1.084***	1.018***
	[8.00]	[10.0]	[11.3]	[11.5]
	(0.027)	(0.027)	(0.027)	(0.025)
Observations (n)	494	494	494	494
R-squared	0.94	0.89	0.87	0.87

Note: Normalized beta coefficients in brackets. Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

of the school resources measured. The results for the share of the budget spent on instructional salaries per student (avg_isal) remained positive and statistically significant for all sixteen models estimated, with a relatively larger beta than the pupil-teacher ratio.

Total expenditures prior to transportation expense (avg_totexp_ntr) explained as much variation in student achievement as the other school variables with beta coefficients ranging from .15 to .35 standard deviations of the dependent variable. The negative sign on this estimate might be explained by the higher expenditures necessary in urban

school districts and the high correlation with instructional salaries.

A primary focus for this study was to analyze the extent to which school district efficiency explained the observed variation in student achievement. The difference in the explanatory power of the specified model after developing a proxy for district efficiency was analyzed by examining the differences in the R² results for the regressions without a measure for district fixed effects and the regressions that include these measures.¹²

Table 4
Grade 7 Reading Scores Post-Estimation WLS Regression Results

Variables	School Year			
	2001–2002	2002–2003	2003–2004	2004–2005
	(1)	(2)	(3)	(4)
pctenroll	-0.533***	-0.605***	-0.594***	-0.568***
	[-0.73]	[-0.79]	[-0.82]	[-0.91]
	(0.016)	(0.015)	(0.020)	(0.015)
avg_t_sal	-0.000000741	-0.000000445	-0.000000516	-0.000000446
	[-0.036]	[-0.022]	[-0.027]	[-0.027]
	(0.00000063)	(0.00000059)	(0.00000067)	(0.00000052)
avg_p_tchr	-0.00795***	-0.0104***	-0.00482***	-0.00774***
	[-0.13]	[-0.19]	[-0.10]	[-0.19]
	(0.0014)	(0.0013)	(0.0016)	(0.0013)
avg_isal	0.0000486***	0.0000582***	0.0000713***	0.0000512***
	[0.18]	[0.21]	[0.28]	[0.24]
	(0.000011)	(0.0000095)	(0.000011)	(0.0000090)
avg_totexp_ntr	-0.0000356***	-0.0000391***	-0.0000312***	-0.0000313***
	[-0.25]	[-0.28]	[-0.26]	[-0.32]
	(0.0000044)	(0.0000040)	(0.0000040)	(0.0000035)
avg_resid	0.946***	0.889***	0.984***	0.887***
	[0.57]	[0.51]	[0.59]	[0.61]
	(0.034)	(0.035)	(0.041)	(0.038)
Constant	1.037***	1.115***	0.820***	1.063***
	[8.22]	[8.41]	[6.50]	[9.70]
	(0.034)	(0.029)	(0.035)	(0.032)
Observations (n)	494	494	494	494
R-squared	0.92	0.93	0.89	0.91

Note: Normalized beta coefficients in brackets. Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 6 shows that after the inclusion of a proxy for district effect the explanatory power of the estimated model increases by fifteen percentage points. The difference in explanatory power remained consistent across all four years in this study. This finding is an important consideration for any measure of school performance or accountability policy. In the absence of a direct measure for district effect, school accountability guidelines may actually only measure student charac-

teristics and the distribution of property wealth given the power of these variables to explain student achievement.¹³ The knowledge of what portion of the variation of student achievement is associated with unobserved district effects combined with the estimates that indicate both the direction and magnitude (Tables 2-5) of that effect, offers a good theoretical foundation upon which to build a school district accountability policy.

Table 5
Grade 8 Math Scores Post-Estimation WLS Regression Results

Variables	School Year			
	2001–2002	2002–2003	2003–2004	2004–2005
	(1)	(2)	(3)	(4)
pctenroll	-0.634***	-0.668***	-0.641***	-0.672***
	[-0.75]	[-0.76]	[-0.78]	[-0.87]
	(0.017)	(0.020)	(0.021)	(0.017)
avg_t_sal	-0.000000121	-0.000000638	-0.000000426	-0.000000119*
	[-0.0052]	[-0.027]	[-0.020]	[-0.059]
	(0.00000068)	(0.00000075)	(0.00000073)	(0.00000061)
avg_p_tchr	-0.0123***	-0.00937***	-0.0121***	-0.00793***
	[-0.18]	[-0.14]	[-0.23]	[-0.16]
	(0.0015)	(0.0017)	(0.0017)	(0.0015)
avg_isal	0.0000647***	0.0000528***	0.0000569***	0.0000814***
	[0.21]	[0.17]	[0.20]	[0.31]
	(0.000012)	(0.000012)	(0.000012)	(0.000011)
avg_totexp_ntr	-0.0000392***	-0.0000323***	-0.0000441***	-0.0000438***
	[-0.24]	[-0.20]	[-0.33]	[-0.36]
	(0.0000047)	(0.0000051)	(0.0000044)	(0.0000041)
avg_resid	1.039***	1.119***	1.092***	1.082***
	[0.54]	[0.55]	[0.58]	[0.60]
	(0.037)	(0.044)	(0.045)	(0.044)
Constant	1.112***	1.010***	1.132***	0.982***
	[7.67]	[6.60]	[7.96]	[7.23]
	(0.037)	(0.037)	(0.038)	(0.037)
Observations (n)	494	494	494	494
R-squared	0.93	0.91	0.91	0.91

Note: Normalized beta coefficients in brackets. Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

In addition, this procedure supplies an objective measure for use in assuring the public and political decisionmakers that funding school districts based on adequacy does not simply reward inefficiency. The objective measurement of district effects provides the means for adjusting legitimate, educationally based, funding differences among districts for the excess costs they encounter due to their own inefficiency.

It is also apparent from Table 6 that district efficiency explains a larger share of the variance in student achievement for the fourth grade than for either the seventh or eighth grades. The fourth grade change is larger for math than for reading. The differences between math and reading narrow in the higher grades. Unobserved effects, for example, school culture, communication, goal orientation, and focus might be more highly associated with early student achievement more than in later grades.

Table 6
Increased Explanatory Power from District Fixed Effects: R-squared Differences

Table 6.1 R-squared for Preliminary Tests on Reading and Math					
	School Year				Average
	2001–2002	2002–2003	2003–2004	2004–2005	
read_gr4_sat	0.78	0.70	0.71	0.72	0.73
read_gr7_sat	0.80	0.84	0.75	0.80	0.80
math_gr8_sat	0.82	0.78	0.79	0.80	0.80
Table 6.2 R-squared for Post Estimation Tests on Reading and Math					
	School Year				Average
	2001–2002	2002–2003	2003–2004	2004–2005	
math_gr4_sat	0.90	0.90	0.89	0.87	0.89
read_gr4_sat	0.94	0.89	0.87	0.87	0.89
read_gr7_sat	0.92	0.93	0.89	0.91	0.91
math_gr8_sat	0.93	0.91	0.91	0.91	0.92
Table 6.3 R-squared Differences					
	School Year				Average
	2001–2002	2002–2003	2003–2004	2004–2005	
math_gr4_sat	0.19	0.20	0.22	0.26	0.22
read_gr4_sat	0.16	0.19	0.16	0.15	0.17
read_gr7_sat	0.12	0.09	0.14	0.11	0.12
math_gr8_sat	0.11	0.13	0.12	0.11	0.12
Average R-squared difference	0.15	0.15	0.16	0.16	0.15

One implication of the disparity of the association of district effect with student achievement depending on grade level comes from separately measuring school accountability or adjusting differential funding by grade. This type of adjustment would be more achievable if the data were available to replicate this study for individual school buildings instead of entire districts.

Conclusions, Implications for Policy, and Further Study

The primary purpose of this study was to test a method for measuring Michigan school district efficiency that could be used to modify a future statewide school funding model based on adequacy. The latter would replace the current resource equity finance system. Besides production efficiency, the desired indicator also gauges "x-efficiency."

This concept evaluates organizational and qualitative attributes of districts not readily observed quantitatively.

The foremost consequence of understanding and measuring the effect of Michigan school district efficiency on student achievement comes from its use to modify Michigan school funding. Redistribution of scarce resources always faces political difficulty and public resistance from those who would bear the burden of providing the benefit to others. Admittedly, this renders a change to an adequacy based Michigan school finance formula politically improbable. However, some future political circumstance, similar to the historical pressure for property tax reform, could materialize and grant unanticipated prominence to this presently dormant policy perspective. Some states have only addressed adequacy of school finance due to actual or threatened

litigation, usually arising out of fresh interpretations of their constitutional educational clause. One genuine objection to adequacy comes from the trepidation for rewarding districts experiencing higher costs precipitated at least partially by factors within their control. The reported results from this research lay the groundwork for minimizing this risk. Identifying the variation in student achievement explained by district effects could help limit funding differences to only the higher costs unrelated to district efficiency.

A second policy implication arising from this research comes from its demonstration of the need for better data. Sacrifices were made regarding the unit of analysis and teacher characteristics precipitated by insufficient data. While this comment hardly seems unexpected from a quantitative researcher, it also represents a common problem for educators across the country, including in Michigan. The need for the retention, ready access, and analysis of student data remains acute in most states. Most states do provide paper reports, lagged by several months, to teachers and administrators regarding student test results. Only five states provide advanced information systems for students and teachers plus offer the means to link the two systems.¹⁴

Michigan should not allow charter schools to avoid reporting crucial data through their use of management companies. An argument based on form that a charter school has no salaries to report cannot be sustained in substance. In essence, the management company pays the salaries as agent for the charter school board of control. Although part of the logic behind charters comes from freedom from bureaucracy, this should not be allowed to interfere with the obligation to demand performance for the investment of tax dollars. This quirk needs to be addressed administratively or by legislation. Neither should Michigan allow bargaining groups or any other special interest to politically prevent the matching of student and teacher performance information.

Previous research has demonstrated that class size reduction has positive effects for student achievement.¹⁵ Some studies reveal diminishing effects for smaller classes.¹⁶ Sometimes they report the positive impact of teacher quality, in addition to the class size measure of teacher quantity.¹⁷ Evidence supporting more cost-effective means of producing positive effects on student achievement may explain the current results controlling for district efficiency.¹⁸ Perhaps improvements in teacher quality can be achieved with aggressive financial incentives to recruit the most qualified and talented people. Organizing learning with higher paid instructional managers supervising larger groups of students assisted by less expensive support staff and technology may leverage teacher resources.

In 2005-2006, Michigan began testing students in contiguous years, as required by the NCLB Act, during grades three through eight for math and English language arts. This will provide the opportunity to measure school performance and efficiency using the student achievement gains accomplished in a single year. It also facilitates the use of lagged student achievement measures as an explanatory variable. This helps account for innate ability and student learning prior to the point of collection for the lagged data. A third enrichment grows out of the ability to measure a single school. This of course assumes that the data elements necessary for school level analysis become available. Student level analysis with linkage to specific classrooms and teachers would provide both increased methodological validity and overall credibility. Direct measures of class size and teacher characteristics also represent improvements. Replication would also be possible using a sample of districts, or even schools, where data was collected directly from the

agency and not from the state.

In addition to the need for further quantitative research, only qualitative study will provide the interpretation of what specific attributes differentiate the districts with positive fixed effects from those that prove negative. Well documented, thorough, and repetitive observations and interviews at sites with the highest and lowest magnitude of fixed effect residuals may be necessary.

Guidance for school districts where funding was adjusted downward as a reflection of inefficiency provides a key ingredient to a school funding system based on adequacy. Meaningful direction will depend on the results of the future research, referred to above, that isolates the elements producing both "x-efficiency" and resource efficiency. Clarity regarding these components provides an essential element in creating a financial incentive for improvement. Only cost differences outside of district control should lead to increased funding. Inefficiencies of the district, that increase cost, should not be rewarded.

This study established the relationship between district effects and student achievement. One policy implication includes the adjustment of district funding by some factor representing the district effect on student achievement, in order to avoid rewarding inefficiency. The actual derivation of an adjustment factor for application to Michigan per pupil school funding represents the seeds for future study. This work should address the limitations previously discussed, especially regarding data quality and more complete measures for student achievement. It should also provide detailed guidance regarding the range of choices and qualitative elements of district efficiency.

Regardless of the actual formula chosen, the care, transparency, and thoroughness of the process for its creation and implementation will help determine utility for transitioning to an adequacy-based school finance system in Michigan. The evidence presented here regarding the relationship between district effects and student achievement provides an introductory, but significant, contribution to this Michigan school finance policy arena.

Endnotes

¹ James Phelps and Michael Addonizio, "How Much Do Schools and Districts Matter? A Production Function Approach to School Accountability" *Educational Considerations* 33 (Spring 2006): 51-62, 30.

² Henry M Levin, "Raising School Productivity: An X-Efficiency Approach," *The American Economic Review* 16 (June 1997): 303-11.

³ Harvey Leibenstein, "Allocative Efficiency vs. 'X-Efficiency,'" *The American Economic Review* 56 (June 1966): 392-415.

⁴ Charles Cobb and Paul Douglas, "A Theory of Production," *The American Economic Review* 18 (March 1928): 139-65.

⁵ Phelps and Addonizio, "How Much Do Schools and Districts Matter?"

⁶ The equation attempts to follow that of Phelps and Addonizio as closely as possible given available data. See Phelps and Addonizio, 53.

⁷ Charter schools are public schools with a charter from one of the allowable organizations under state law and operate independently from the traditional local public school board in their jurisdiction. Since many of these schools are actually operated by management companies contracted by their board of control, they do not technically have or report any salary information on Bulletin 1014. Based on

the full time equivalency pupil count data maintained by CEPI, the number of charter schools for each of the years in the panel was 185, 185, 192, and 210 respectively.

⁸ UCLA Academic Technology Services, "Stata Frequently Asked Questions: What Type of Weights Do SAS, Stata, and SPSS Support?" <http://www.ats.ucla.edu/stat/Stata/faq/weights.htm>.

⁹ Tables are omitted in the interest of brevity but are available from the author.

¹⁰ Phelps and Addonizio, "How Much Do Schools and Districts Matter?"

¹¹ Ibid.

¹² This technique follows the analysis of Phelps and Addonizio.

¹³ Ibid.

¹⁴ David J. Hoff, "Technology Counts: Delving Into Data," *Education Week*, 2006 25 (May 4 2006), 12.

¹⁵ Ronald Ferguson, "Paying for Public Education: New Evidence on How and Why Money Matters," *Harvard Journal on Legislation* 28 (May 1991): 465-98; Larry Hedges, Richard Laine, and Rob Greenwald, "Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs and Student Outcomes," *Educational Researcher* 23 (May 1994): 5-14; Barbara Nye, Larry G. Hedges, and Spyros Konstantopoulos, "The Long-Term Effects of Small Class Sizes: A Five-Year Follow-up of the Tennessee Class Size Experiment," *Educational Evaluation and Policy Analysis* 21 (Summer 1999): 127-42; Gary Peevely, Larry Hedges, and Barbara Nye, "The Relationship of Class Size Effects and Teacher Salary," *Journal of Education Finance* 31 (Summer 2005): 101-09.

¹⁶ Ferguson, "Paying for Public Education."

¹⁷ Peevely et al., "The Relationship of Class Size Effects and Teacher Salary."

¹⁸ Lynn Ilon and Anthony H Normore, "Relative Cost Effectiveness of School Resources in Improving Achievement," *Journal of Education Finance* 31 (Winter 2006): 238-54.

Commentary

Unacceptable *but* Tolerated behavior

Anne L. Jefferson

The literature discusses bullying in terms of the misuse of a power situation over another individual repeatedly.¹ Single, isolated incidences do not qualify as an act of bullying. Rather, bullying is the repetition of these acts combined with the desire on the part of the individual with the greater power base to cause physical, emotional, or social distress in another individual. Bullying is not acceptable in a civilized society, and, increasingly, it is recognized as a punishable act. However, the seriousness of bullying is often addressed differently across types of educational organizations.

Within school systems and universities, great pain is taken to develop and enforce policy, guidelines, and procedures on the prevention of the mistreatment of students by other students or staff. If we turn briefly to school systems, we find many schools and school systems with a policy including guidelines and procedures to follow should a student be the subject of bullying. For example, in 2001, the Michigan School Board Association passed an updated policy on bullying and hazing. This policy was later given further clarification by Robert Ebersole, the Assistant Director of Bylaw and Policy Services.² Bullying and hazing were to be considered forms of harassment. In 2004, the Cambridge (Massachusetts) School Committee produced its finalized version of administrative procedures and guidelines on prevention of bullying.³ In 2005, the Victoria (Australia) State Department of Education and Early Childhood Development, reviewed and updated anti-bullying policies and practices in its government schools.⁴ At the university level, the Open University (United Kingdom) has an extensive web site informing students about university policy on bullying and harassment along with procedures to follow and forms to file if they are the subject of such treatment.⁵ Similar policies against student bullying have been adopted by institutes of higher learning across North America, Australia, New Zealand, and a number of the Scandinavian and European countries.

What appears to be less frequently addressed, especially by institutes of higher learning in North America, is administrative bullying, oftentimes referred to as workplace bullying. According to Gary Namie, Co-founder of the Workplace Bullying and Trauma Institute (WBTI), workplace bullying is “deliberate, repeated, health-impairing mistreatment of an employee.”⁶ Although there seems to be a common understanding of the harm caused by student bullying across school (K-12) systems and higher education institutions and the need for institutional protections and actions, there is a noticeable absence of similar policies and procedures when the alleged bully is a higher education administrator. In contrast, one will find policies and procedures related to sexual harassment well-ingrained in higher education, to the extent that a specific office or department is designated as a place to

deal with these offenses. On the other hand, harassment in the form of administrative bullying tends to be very generally attended to. At best, it might be alluded to in a general way in university policy with a statement to the effect that the administration has responsibility to provide a safe and healthy working environment. Missing from such generic statements is an acknowledgement that administrative bullying exists and hence the administration has a responsibility to address it. Through this denial, no further action by the administration is needed, for example, to define workplace bullying, clarify institutional responsibility for addressing complaints, or to provide employees with guidelines and procedures for reporting workplace bullying. In other words, the administration feels no responsibility to provide the same standard of protection for its employees as it does for students. The implication and, too often, reality is the tolerance of unacceptable behavior by one of their own. This unwillingness to self-police opens the door for administrative bullying.

Absent such policies and protections, the administration's typical response to an employee's claim of workplace bullying is to suggest that a “personality clash” exists and the party with the lower power base should look within herself or himself for a solution. Oftentimes, if the bullying or “personality clash” continues, the solution strongly encouraged, directly or indirectly, by the administration and the individual's peers is departure from the working environment, regardless of the potential professional harm and personal disruption this might cause. On the other hand, the bullying administrator rarely suffers any negative consequences and usually remains in a position of authority. Noveck speaks directly to this scenario in her discussion of the “nasty boss phenomenon,” with a quote from Jeffrey Pfeffer, professor at Stanford's Graduate School of Business, that is very revealing: “Certainly, the behavior of nasty bosses is way more public than it used to be But does it have consequences? I just don't see it.”⁷

The lack of negative consequence for administrators who abuse their power through bullying employees is detrimental not only to the person(s) being bullied but also to the organization that tolerates it. For example, Finkelstein identified staff departures and high turnover as potential consequences of administrators who “ruthlessly eliminate” underlings who do not give them total and unquestioning support,⁸ a common type of administrator bullying. While Finkelstein was referring to CEO's bullying of employees in the private sector, academia is similarly fertile ground for administrator bullying of faculty members, particularly, but not exclusively, newly hired academics or assistant professors. Given their long probationary period,⁹ assistant professors may be at greater risk of being bullied. This academic tradition essentially enables the bullying administrator to more easily identify potential targets.

The administrator's ability to get away with bullying rests upon inequalities in power, the lack of institutional safeguards for those who might become targets of the bully, and the lack of sanctions which serve as punishment and deterrent for bullying. In the absence of institutional safeguards and sanctions, a faculty member who makes a claim of bullying against an administrator risks becoming the subject of administrative scrutiny, rather than vice versa. As part of the institution's administration, the bully may well be given a shield of protection and even provided with free legal advice and assistance from university counsel as though he or she were the target rather than the perpetrator. At the same time, it is unlikely that the faculty member, although an employee of the institution just like the administrator, will enjoy these benefits. Retaliatory sanctions against the faculty

Anne L. Jefferson is Professor of Education Finance and Administration at the University of Ottawa.

member, such as being reprimanded by the administration for initiating a “false accusation” and being warned (threatened) that another such “false accusation” might result in more severe administrative sanctions, are not uncommon. In such cases, the faculty member, not the administrator, is called “on the carpet” for daring to voice objections to being bullied.

These actions by the administration serve to silence the faculty member and embolden the bullying administrator. In institutions where faculty are unionized, one could legitimately ask, where is the faculty union under such circumstances? Unfortunately, many an academic union views itself as powerless to act against administrator bullying. In cases where there is no institutional infrastructure to address administrative bullying, the unions’ only instrument in dealing with it is through the collective agreement. If the collective agreement is silent on this issue, the faculty member can expect little union support. The administrator is now free to escalate bullying behavior and act with impunity, ignoring normal protections faculty take for granted. If the faculty member protests, the bullying administrator may now label her or him a “troublemaker” who is interfering with the work of the Faculty.

A potential consequence or byproduct of administrative bullying, e.g., where the bully refers to the faculty member as a “troublemaker” in the presence of other faculty and by doing so encourages group bullying, is “mobbing.”¹⁰ Leymann describes mobbing as a “nonviolent, polite, sophisticated” approach to bullying by a group of coworkers in “ostensibly rational workplaces” and noted: “Universities are an archetype.”¹¹ In universities, mobbing behavior may, in the initial stages, take the form of “wear(ing) the target down emotionally by shunning, gossip, ridicule, bureaucratic hassles, and withholding of deserved rewards.”¹² Mobbing behavior may escalate to “formal outbursts of aggression” whereby “some real or imagined behavior” is asserted as “proof of the target’s unworthiness to continue in the normal give-and-take of academic life.”¹³ At the initial stages, the administrative bully may simply stand on the sidelines and encourage mobbing, but as it escalates the bully may use it as an opportunity to invoke or threaten to invoke disciplinary measures against the faculty member without establishment of the facts. The administrative bully may even make formal charges of “misconduct” where false charges against the faculty member are aired at higher levels of university administration or in front of a campus tribunal. Westhues refers to these events as “degradation rituals” which leave the faculty member with two stark and unpleasant options: quit or fight for their professional rights and life.¹⁴

As mentioned previously, administrative bullying of faculty is not limited to assistant professors. Uscilka described the case of Bill Lepowsky, a professor with 37 years experience at a college, who was falsely accused by an administrator of “violating procedures related to textbook adoption, textbook printing, and textbook sales to students. ... accused of saying and doing things ..., threatened with termination, and denied a sabbatical.”¹⁵ Although the college never undertook a full investigation, Lepowsky was eventually able to clear his name with the assistance of colleagues and the faculty union, and ultimately he received an apology from the college chancellor. Even so, the bullying continued for another year, and only after a change in the administration did the abuse finally stop.

Elash stated, “Even if they are well intentioned, leaders can abuse their power. ... Some are just bullies who mistreat others simply because they are in a position to do so.”¹⁶ The administrator’s claim is I am just tough and demanding, and look how much more profitable the

organization is. The bottom line becomes the justification, but the bottom line has a number of interpretations. In the world of academia, the bottom line is the creation and advancement of knowledge through highly educated and skilled faculty. The traditional division of authority between labor and management in the private sector is often less clear between faculty and administration in higher education institutions. The insecurities and weaknesses of an administrator, especially one who is trying unsuccessfully to bridge academic and managerial expectations, are perhaps more open for display, discussion, and even challenge by faculty. These types of administrators may be more likely to engage in bullying and harassment in an attempt, for example, to deflect attention from their own shortcomings or to spite those who are more successful. Without consequences, unacceptable behavior becomes part of the norm. The norm is what has been agreed to, not formally but by practice, as tolerable behavior.

Endnotes

¹ See, for example, *Literature Review: Selected References* (Toronto, Ontario: Ontario Public School Boards’ Association) http://www.opsba.org/Policy_Program/Interesting_Programs/Bullying/literature_review.html.

² Robert Ebersole, *Bullying and Hazing Policies Offered* (Lansing, MI: Michigan Association of School Boards) www.masb.org/pdf/bullyhazpolicy.pdf.

³ Cambridge Public Schools, *Cambridge Public Schools Administrative Procedures and Guidelines on Prevention of Bullying* (Cambridge, MA) www.cpsd.us/Web/PubInfo/Anti-bullying_procedures.pdf.

⁴ State of Victoria, *Safe Schools Are Effective Schools* (Victoria, Australia: Department of Education and Early Childhood Development) <http://www.sofweb.vic.edu.au/wellbeing/safeschools/bullying>.

⁵ The Open University, *Bullying and Harassment: Student Code for Dealing with Bullying and Harassment*, <http://www3.open.ac.uk/our-student-policies/pdf/bullying.pdf>.

⁶ Gary Namie as quoted in “Laney (College) Hosts ‘Workplace Bullying’: Trustee Handy encourages employees to speak out,” by Mary Uscilka, *The Laney Tower*, May 11, 2006, <http://www.bullyinginstitute.org/press/laney051106.html>.

⁷ Jocelyn Noveck, “The Nasty Boss Phenomenon,” *The Ottawa Citizen*, July 6, 2006, D3.

⁸ Sidney Finkelstein, cited in “Seven deadly habits of CEO’s,” by Ray Williams, *National Post*, June 14, 2006, FP8.

⁹ The normal probation period for an assistant professor is six years. Although this probationary period varies and can be shorter, most academics would not receive consideration for tenure earlier than two years.

¹⁰ Ken Westhues, “The Unkindly Art of Mobbing,” *Academic Matters: The Journal of Higher Education* (Fall 2006): 18-19.

¹¹ Leymann cited in Westhues, 18

¹² Westhues, 18.

¹³ Ibid., 19.

¹⁴ Ibid., 19.

¹⁵ Uscilka, “Laney (College) Hosts ‘Workplace Bullying.’”

¹⁶ Daniel Elash, *When the Boss Is a Bully*,
<http://www.leader-values.com/content/detail.asp?ContentDetailID=338>.

Subscribe TODAY!

to

educational considerations

Educational Considerations is a leading peer-reviewed journal in the field of educational leadership.

Educational Considerations is published twice yearly by the College of Education at Kansas State University.

Educational Considerations invites subscribers for only \$13.00 annually. Subscribers receive paper copy and electronic copy.

OR

Save 20% on the regular subscription price when you select electronic copy only!

***** ORDER FORM *****

Please send me:

- ☐ Paper copy \$13.00 (electronic copy included at no extra cost) for one year subscription
- ☐ Electronic copy only \$10.40 for one year subscription

Name _____

Address _____

City _____ State _____ Zip _____

Make checks payable to *Educational Considerations*.

Mail with order form to:

Editor, Educational Considerations, Bluemont Hall, 1100 Mid-Campus Drive,
Kansas State University, Manhattan, KS 66506

Visit Us Online at:
<http://coe.ksu.edu/EdConsiderations/>

ISSUES 1990-2007

Educational Considerations is a leading peer-reviewed journal in the field of educational leadership. Since 1990, *Educational Considerations* has featured outstanding themes and authors relating to leadership:

SPRING 1990: a theme issue devoted to public school funding.

Edited by David C. Thompson, Codirector of the UCEA Center for Education Finance at Kansas State University and Board of Editors of *Educational Considerations*.

FALL 1990: a theme issue devoted to academic success of African-American students.

Guest-edited by Robbie Steward, University of Kansas.

SPRING 1991: a theme issue devoted to school improvement.

Guest-edited by Thomas Wicks & Gerald Bailey, Kansas State University.

FALL 1991: a theme issue devoted to school choice.

Guest-edited by Julie Underwood, University of Wisconsin-Madison and member of the Editorial Advisory Board of *Educational Considerations*.

SPRING 1992: a general issue devoted to philosophers on the foundations of education.

FALL 1992: a general issue devoted to administration.

SPRING 1993: a general issue devoted to administration.

FALL 1993: a theme issue devoted to special education funding.

Guest-edited by Patricia Anthony, University of Massachusetts-Amherst and member of the Editorial Advisory Board of *Educational Considerations*

SPRING 1994: a theme issue devoted to analysis of funding education.

Guest-edited by R. Craig Wood, Codirector of the UCEA Center for Education Finance at the University of Florida and member of the Editorial Advisory Board of *Educational Considerations*.

FALL 1994: a theme issue devoted to analysis of the federal role in education funding.

Guest-edited by Deborah Versteegen, University of Virginia and member Editorial Advisory Board of *Educational Considerations*.

SPRING 1995: a theme issue devoted to topics affecting women as educational leaders.

Guest-edited by Trudy Campbell, Kansas State University.

FALL 1995: a general issue devoted to administration.

SPRING 1996: a theme issue devoted to topics of technology innovation.

Guest-edited by Gerald D. Bailey and Tweed Ross, Kansas State University.

FALL 1996: a general issue of submitted and invited manuscripts on education topics.

SPRING 1997: a theme issue devoted to foundations and philosophy of education.

FALL 1997: first issue of a companion theme set (Fall/Spring) on the state-of-the-states reports on public school funding.

Guest-edited by R. Craig Wood (University of Florida) and David C. Thompson (Kansas State University).

SPRING 1998: second issue of a companion theme set (Fall/Spring) on the state-of-the-states reports on public school funding.

Guest-edited by R. Craig Wood (University of Florida) and David C. Thompson (Kansas State University).

FALL 1998: a general issue on education-related topics.

SPRING 1999: a theme issue devoted to ESL and Culturally and Linguistically Diverse populations.

Guest edited by Kevin Murry and Socorro Herrera, Kansas State University.

FALL 1999: a theme issue devoted to technology.

Guest-edited by Tweed Ross, Kansas State University.

SPRING 2000: a general issue on education-related topics.

FALL 2000: a theme issue on 21st century topics in school funding.

Guest edited by Faith Crampton, Senior Research Associate, NEA, Washington, D.C.

SPRING 2001: a general issue on education topics.

FALL 2001: a general issue on education funding.

SPRING 2002: a general issue on education-related topics.

FALL 2002: a theme issue on critical issues in higher education finance and policy.

Guest edited by Marilyn A. Hirth, Purdue University.

SPRING 2003: a theme issue on meaningful accountability and educational reform.

Guest edited by Cynthia J. Reed, Auburn University, and Van Dempsey, West Virginia University.

ISSUES 1990-2007 continued

FALL 2003: a theme issue on issues impacting on higher education at the beginning of the 21st century.
Guest edited by Mary P. McKeown-Moak, MGT Consulting Group, Austin, Texas.

SPRING 2004: a general issue on education topics.

FALL 2004: a theme issue on issues relating to adequacy in school finance.
Guest edited by Deborah A. Verstegen, University of Virginia.

SPRING 2005: a theme issue on reform of educational leadership preparation programs.
Guest edited by Michelle D. Young, University of Missouri; Meredith Mountford, Florida Atlantic University; and Gary M. Crow, The University of Utah.

FALL 2005: a theme issue on reform of educational leadership preparation programs.
Guest edited by Teresa Northern Miller, Kansas State University.

SPRING 2006: a theme issue on reform of educational leadership preparation programs.
Guest edited by Teresa Northern Miller, Kansas State University.

FALL 2006: a theme issue on the value of exceptional ethnic minority voices.
Guest edited by Festus E. Obiakor, University of Wisconsin-Milwaukee.

SPRING 2007: a theme issue on educators with disabilities.
Guest edited by Clayton E. Keller, Metro Educational Cooperative Service Unit, Minneapolis, Minnesota, and Barbara L. Brock, Creighton University.

FALL 2007: a theme issue on multicultural adult education.
Guest edited by Jeff Zacharakis and Gabriela Díaz de Sabatés, Kansas State University, and Dianne Glass, Kansas Department of Education.