9-1998

An Analysis of 4th and 5th Grade MEAP Results at Elementary Schools in Southwest Michigan

Kevin M. Hollenbeck
W.E. Upjohn Institute for Employment Research, hollenbeck@upjohn.org

Kristine Kracker
W.E. Upjohn Institute for Employment Research

An Analysis of 4th and 5th Grade MEAP Results at Elementary Schools in Southwest Michigan

Kevin Hollenbeck
Kristine Kracker

September 1998

1. Purpose of the Study

The purpose of this study is to statistically examine 4th and 5th grade standardized test (MEAP) results from buildings located in Kalamazoo County and its surrounding counties. This empirical work is being conducted for two reasons. First, we want to analyze the relationship between a building's performance on the tests and the economic circumstances of its students. It is generally believed that there is a positive or direct relationship between student achievement as measured by test results and the household income levels of students, i.e., schools that have students from higher income backgrounds will have better test results. This study examines test score results of elementary schools from districts throughout southwest Michigan to see whether that relationship holds and, if so, to get a sense of the strength and shape of that relationship.

The second reason for examining test results at the building level is to determine whether there are schools in Kalamazoo County or other parts of southwest Michigan that perform much better or much worse than might be expected by the economic circumstances of its students. If it were the case that there is substantial variability in the relationship between income and test results, then it would be inappropriate to make a judgement about the school's educational effectiveness based on the income or poverty levels of its students. In this case, students from a school that has a high percentage of its students coming from poor backgrounds may perform better than students from a school with a much more advantaged student body.

Many factors explain a building's test results: background characteristics of the students and their families, the curriculum and instruction that the students have been exposed to, effectiveness of the teachers that the students have had,[1] characteristics of the building such as class size and the building's staff, and characteristics of the district. In other words, an individual student's test score

[1]Note that the teacher effect operates through **all** teachers that students have been exposed to not just their current teacher.

is a function of all of these factors. Any attempt to analyze students' test scores should take all of these factors into account. Unfortunately, we systematically measure very few of these characteristics, so our attempts to explain statistically test results can be partial at best.

Among all of the factors that influence test scores, this study focuses on family income levels. The underlying hypothesis is that the family income of students in a building is strongly and positively related to test performance of students. Holding everything else constant–teacher qualities, building characteristics, district characteristics, and other background characteristics of the student–a higher level of average family income will increase the percentage of students who satisfactorily pass the test. Another way of stating this hypothesis is that there is a negative relationship between poverty levels and test performance.

Several theoretical reasons underlie this hypothesis. First, higher levels of family income indicate that students' families have more resources to spend on their children's education. Poverty forces scarcity onto families who must carefully use their resources, and therefore they may not have much to spend toward education. Second, if we accept that there is a genetic component to intelligence that is positively related to test scores and educational achievement, then it may be the case that this same genetic component is related to family income. A positive relationship between family income and student test scores may be caused by parental intelligence being positively related to family income and to child's intelligence and test scores. Third, there may be peer effects that are important in learning. A student will do better (do worse) on a test if he or she has more (or fewer) classmates who do well on the test.

This central hypothesis is not new or surprising; many studies have confirmed it. It does imply that policymakers and administrators who want to reduce performance differentials between buildings may have a more difficult task if there are significant differentials in the economic circumstances of the buildings' students. But the variability in test outcomes across buildings that have students with similar economic backgrounds leads to a secondary hypothesis. That hypothesis states that school buildings that have test results that exceed most other buildings with similar levels of student family income are more effective.[2] The study examines whether school buildings in Kalamazoo County perform very well or, conversely, very poorly given their students' backgrounds.

The next section of the paper presents a discussion of the data and statistical methodology used to test the central hypotheses. That section is followed by a presentation of the results of the statistical estimates.

[2]It could be the case that there are non-school related reasons for test performance that is far above or far below what would be expected based on the poverty level of the building's students. However, it seems far more likely that atypical results are caused by the educational practices (effectiveness) of the school and its teachers.

2. Methods and Data

The Michigan Department of Education (MDE) has placed considerable data about districts, buildings, and MEAP results on the Internet so that they are easily accessible to the general public. We downloaded these data for the 18 counties that comprise southwest Michigan (roughly the corner of the State bounded by Lake Michigan on the west, the state line on the south, a line connecting Jackson to Lansing on the east, and a line connecting Lansing to Muskegon on the north). The specific counties are as follows:

Allegan	Cass	Ionia	Muskegon
Barry	Clinton	Jackson	Ottawa
Berrien	Eaton	Kalamazoo	St. Joseph
Branch	Hillsdale	Kent	Van Buren
Calhoun	Ingham		

The statistical analysis examines 4th grade reading and math test data and 5th grade science and writing data from the 1994/95 school year to the 1997/98 school year. We chose to examine elementary school test data for three reasons. First, we believe that the free- or reduced-price lunch eligibility data are more accurate at that level than in higher grades, where there is severe underreporting. Second, there are far more elementary buildings than middle schools or high schools, so that the sample for analyses would be larger. Third, there is substantial evidence that a student's academic performance in secondary school is highly correlated with his or her performance in elementary school, so the results of our analyses generalize beyond the elementary level. We started with 1994/95 data since it was the first year after Proposal A, so all observations would be comparable with that respect.[3]

[3]We estimated models using 1993/94 data, which is the earliest years that are easily accessible through the Internet, and found that the results were very similar to the post-Proposal A results.

The unit of analysis is the building. The dependent variable is the percentage of students who scored in the satisfactory range or who were classified as proficient on the MEAP tests. Note that this variable only pertains to grades that took the test–4th grade for reading and math; 5th grade for science and writing. The explanatory (or independent) variables are the percentage of students eligible for free- or reduced-price lunches and the building's administrative and instructional expenditure per student.[4] The statistical technique that has been used is multiple regression, which is essentially trying to find the best line that fits the data. The following equation was estimated:

$$\text{(1)} \quad MEAP_{it} = a + b\ FRL_{it} + c\ EXP_{it} + e_{it}$$

where $MEAP_{it}$ =	percentage of students in building i who got a satisfactory (proficient) rating on the MEAP in year t
FRL_{it} =	percentage of students in building i who are eligible for free- or reduced-price lunches in year t
EXP_{it} =	administrative and instructional expenditures per student in building i in year t
e_{it} =	error term
a, b, c =	estimated coefficients

The "b" coefficient that is estimated is interpretable. It is the slope coefficient of the relationship between poverty and test results, or in other words, quantifies the marginal effect of students in poverty. If we plot test score on the vertical axis of a graph and free- or reduced-price lunch percentage on the horizontal axis, then the slope coefficient (assuming a negative relationship) indicates the *predicted* decrease in the percentage of students with a satisfactory score that occurs

[4]We did considerable testing of different models, some of which included other building-level information and district-level information as independent variables. We chose the simple model with two independent variables for several reasons–the other variables were not statistically related to test scores, the other variables could not be easily interpreted or were not theoretically important, district-level variables had biased standard errors because of aggregation, or for parsimony.

when there is a percentage increase in students in poverty. So, for example, if the slope coefficient were -.75, then a 10 percentage point increase in poverty would be predicted to result in a 7.5 percentage point decrease in the percent of students with a satisfactory score.

We would expect the slope coefficient (in absolute value) to lie between 0 and 1. We may interpret this number as indicative of how well elementary schools are able to "reach" disadvantaged youth. If the number were nearly 0, then the buildings do relatively well with children in poverty. The predicted overall passing rate would not change much, on average, when the percentage of disadvantaged students increases. If the number were close to 1, then the percentage of students who are eligible for free- or reduced-price lunches would seem to have a strong influence on the test results. For the average building, a one percentage point increase in the poverty rate of students would decrease the predicted passing rate of all students by nearly one percentage point in this case.[5]

In equation (1), the mathematical expression "$a + c * EXP_{it}$" represents a prediction of what the typical building's passing rate would be if it had no students eligible for free- or reduced-price lunches given the its average expenditure per pupil. In general, this expression is between 60 to 80 percent.

For each of the four MEAP tests, we experimented with the functional form of the regression model to find whether or not there was a linear or curvilinear relationship between poverty (as measured by a building's free- or reduced-price lunch percentage) and test scores. In other words,

---

[5]Note that we are not predicting the passing rate of students who are or are not eligible for free- or reduced-price lunches. The slope coefficient is derived by analyzing the results from different buildings with different students, who have different poverty rates.

we wanted to know whether the line that best fit the data was a straight line, a quadratic (parabolic), a cubic (two inflection points), or a quartic (three inflection points).[6]

In regression analysis, one measure of how well the statistical model performs is called R-squared. This statistic is the percentage of the variance in the dependent variable that is explained by the independent variables in the model. R-squared statistics range from 0 to 1. If there is a "perfect" or very good relationship between the explanatory variables–free- or reduced-price lunch eligibility and expenditures per pupil–and the dependent variable–MEAP score–then the R-squared statistic will be close to 1. That means that the variation that occurs in the explanatory variables exactly matches the variation in the dependent variable. If the independent variables are not related to the dependent variable at all, the R-squared statistic will be 0. The models that are presented in this study have R-squared's of around .50, meaning that the variables that we have used explain about 50 percent of the test score results and the missing variables that we don't include in the model(or cannot observe) also explain about 50 percent.

[6]Specifically, we estimated the following equations in addition to (1):

$$MEAP_{it} = a + b_1\, FRL_{it} + b_2\, (FRL_{it})^2 + c\, EXP_{it} + e_{it} \tag{2}$$

$$MEAP_{it} = a + b_1\, FRL_{it} + b_2\, (FRL_{it})^2 + b_3\, (FRL)^3 + c\, EXP_{it} + e_{it} \tag{3}$$

$$MEAP_{it} = a + b_1\, FRL_{it} + b_2\, (FRL_{it})^2 + b_3\, (FRL)^3 + b_4\, (FRL)^4 + c\, EXP_{it} + e_{it} \tag{4}$$

3. Results

Table 1 presents the regression estimates for all four MEAP tests. Note that two sets of results are presented for each test. The top panel of the table presents results from the first set of regressions that use average passing rates for a building over the last 3 - 4 years (four-year averages for reading, math, and science; three-year average for writing) and the three-year average free- or reduced-price lunch percentage over the years 1994/95 through 1996/97. The advantage of this set of results is that it averages out any unique or idiosyncratic results that might have occurred in a given year, when there was an unusually high or low percentage of students in poverty who were in

Table 1
Regression Estimates

Independent Variable	4th Grade Reading	4th Grade Math	5th Grade Science	5th Grade Writing
Panel 1: Variables are averaged				
Constant	65.38	81.68	59.65	75.24
Average per pupil expenditure ($000)	.0022	--	--	.0023
% Free/reduced-price lunch	-1.17	-.48	-.44	-.82
$(\%\text{Free/reduced-price lunch})^2$	.015	--	--	.009
$(\%\text{ Free/reduced-price lunch})^3$	-.000087	--	--	-.000056
$\overline{R}^2$	.5019	.4971	.5195	.4070
Sample size	483	483	432	432
Panel 2: 1997/98 MEAP tests				
Constant	69.92	88.30	56.19	81.41
Average per pupil expenditure ($000)	.0024	--	--	--
% Free/reduced-price lunch	-.70	-.40	-.49	-.59
$(\%\text{ Free/reduced-price lunch})^2$	.0024	--	--	.0020
$(\%\text{ Free/reduced-price lunch})^3$	--	--	--	--
$\overline{R}^2$	.3761	.3167	.3705	.2765
Sample size	483	483	432	432

-- Not in final model.

the test-taking classes. The second set of results, presented in the bottom panel of table 1, uses 1997-98 test results and 1996-97 free- or reduced-price lunch percentage. This set of results uses the most recent data available. The disadvantage to this set of results is that it might be idiosyncratic, but on the other hand, school districts or buildings may have changed educational practices in response to their past test performances, so this set of results would reflect current practice. We discuss each of the test results in the following sections.

Reading. The best-fitting curve for the reading test results when we estimated the relationship using four-year building averages as the dependent variable was a cubic that included average expenditures per student (see panel 1 of table 1). Figure 1 provides a graph of the curve and a scatter diagram of all of the buildings in the data set. Notice that the slope of the curve goes through three ranges–it is relatively steep up to a free or reduced price lunch percentage of about 30

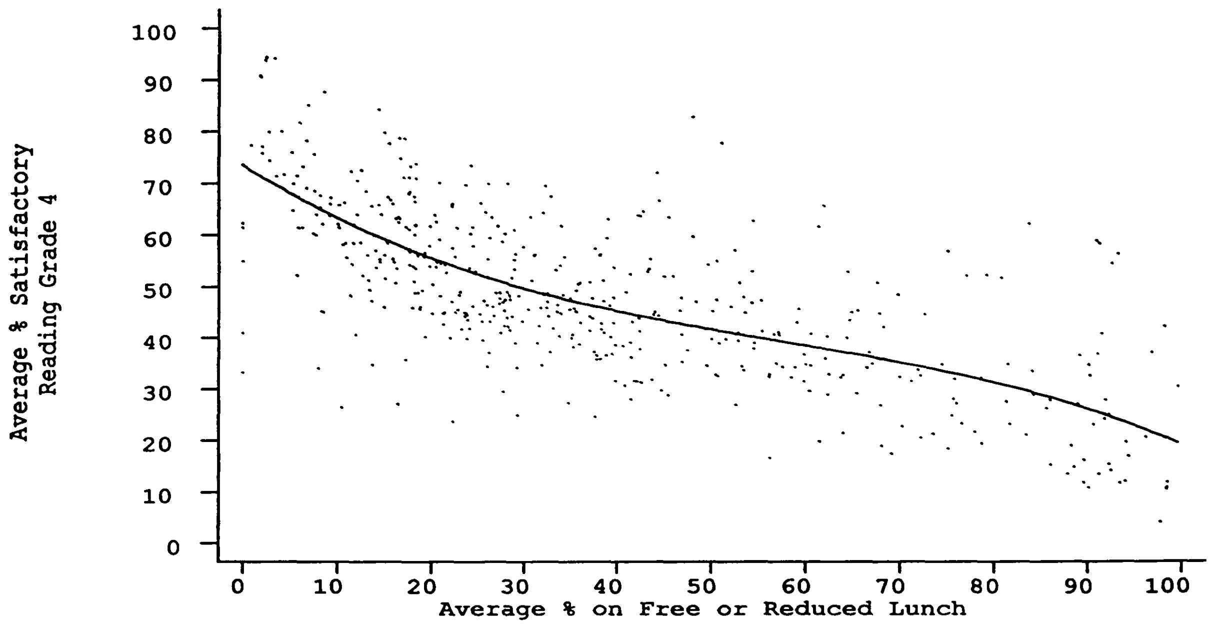

Figure 1. 4th Grade Reading Results Using Averaged Data.

percent, it is then relatively flat up to about 70 percent, and then it gets relatively steep again. In fact, the slopes at 10 percent, 50 percent, and 90 percent are -.897, -.330, and -.595, respectively.

To interpret these slopes, recall that the y-intercept shows the passing rate for the average building as if there were no students who were eligible for free- or reduced-price lunches. That y-intercept is 73.62 percent.[7] With a slope coefficient of -.897 when the free- or reduced-price lunch percentage is 10 percent, the regression curve indicates that the typical building's passing rate is quite sensitive to its percentage of disadvantaged students. A one-percentage point increase in the free- or reduced-price lunch eligibility rate reduces the overall the passing rate of the reading test by .897 percentage points for buildings with an eligibility percentage of around 10 percent. For school buildings that have a free- or reduced-price lunch percentage of 50 percent, the slope coefficient of -.330 indicates that passing rates are less sensitive to the poverty rate. For the "typical" buildings in that range, a one percentage point increase in the free- or reduced-price eligibility rate reduces the overall passing rate by .33 percentage points. Finally, if the free- or reduced-price lunch percentage is 90 percent, the slope coefficient of -.595 indicates that the predicted decrease in the passing rate for the typical school resulting from a one percentage point increase in the poverty rate would be .595 percentage points.

The scatter plot in figure 1 displays considerable variability in the data. Many points are far from the regression curve. Notice that there are schools with average free- or reduced-price lunch percentages near 0 that have average passing percentages of less than 50 percent. Others are over 90 percent. Similarly, there are buildings with free-or reduced-price lunch percentages of around 50 percent that are near the top in terms of average passing rates–80 percent or more.

---

[7]The slope coefficient is 65.38, the coefficient on expenditures per pupil is .0022, and the mean expenditure per pupil is \$3744. The y-intercept is then 65.38 + .0022 * 3744 = 73.62 .

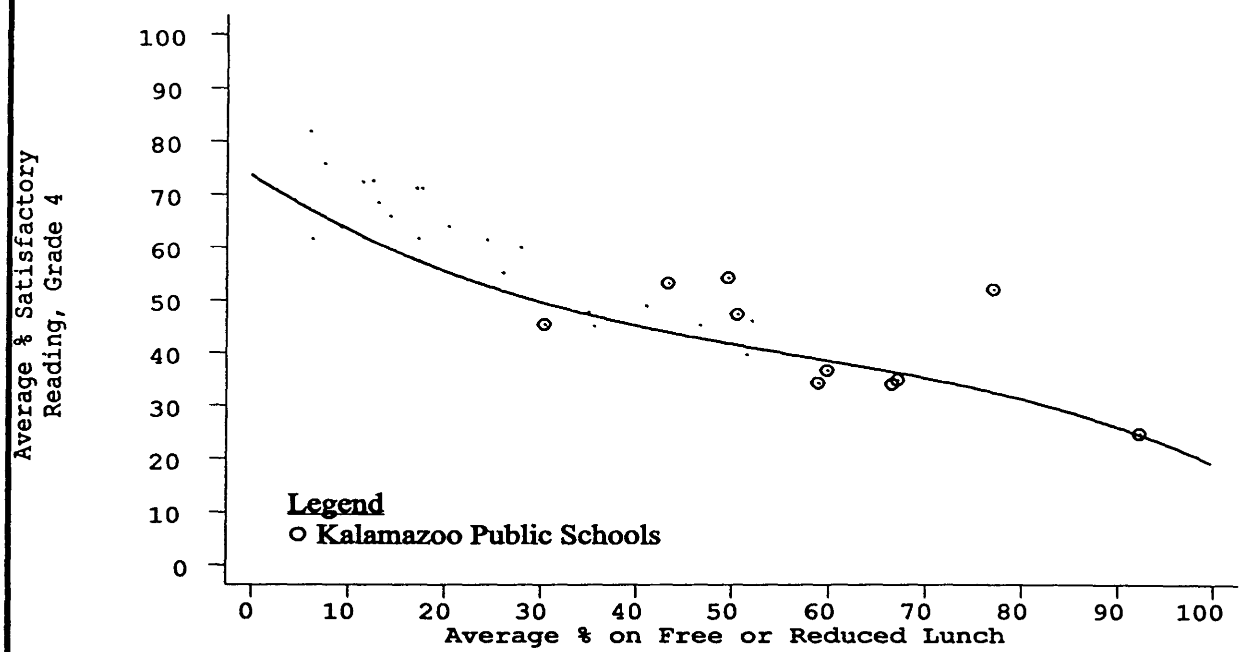

Figure 2. Kalamazoo County 4th Grade Reading Results Using Averaged Data

Figure 2 shows the best-fitting curve with elementary buildings from only Kalamazoo County displayed in the scatter diagram. The points that are circled represent Kalamazoo Public Schools (KPS) buildings. Most of the points in figure 2 are "close" to the regression curve. The amount of statistical uncertainty around each point in the data depends on the data values, and is unique for each point. In general, a building must be about 20 percentage points high or low to be considered as performing above or below average, with statistical significance and none of the Kalamazoo County schools are that far from the regression curve.

Figures 3 and 4 display the best regression line when only the most recent year of reading test data are analyzed. In this case, the best curve is a quadratic. Again, the slope coefficient changes throughout the curve, but in this case, the coefficient gets smaller and smaller as the free- or reduced lunch eligibility percentage increases. At 10 percent free- or reduced-price lunch eligibility, the coefficient is -.650; at 50 percent, it is -.461; and at 90 percent, it is -.271. In this

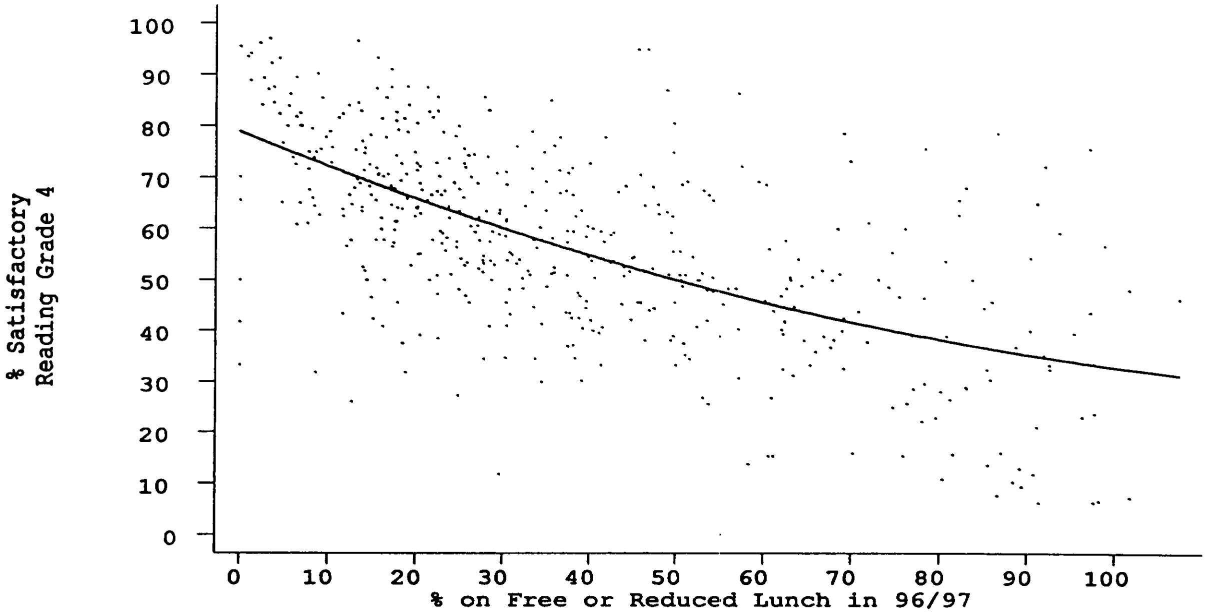

Figure 3. 4th Grade Reading Results Using Most Recent Year's Data.

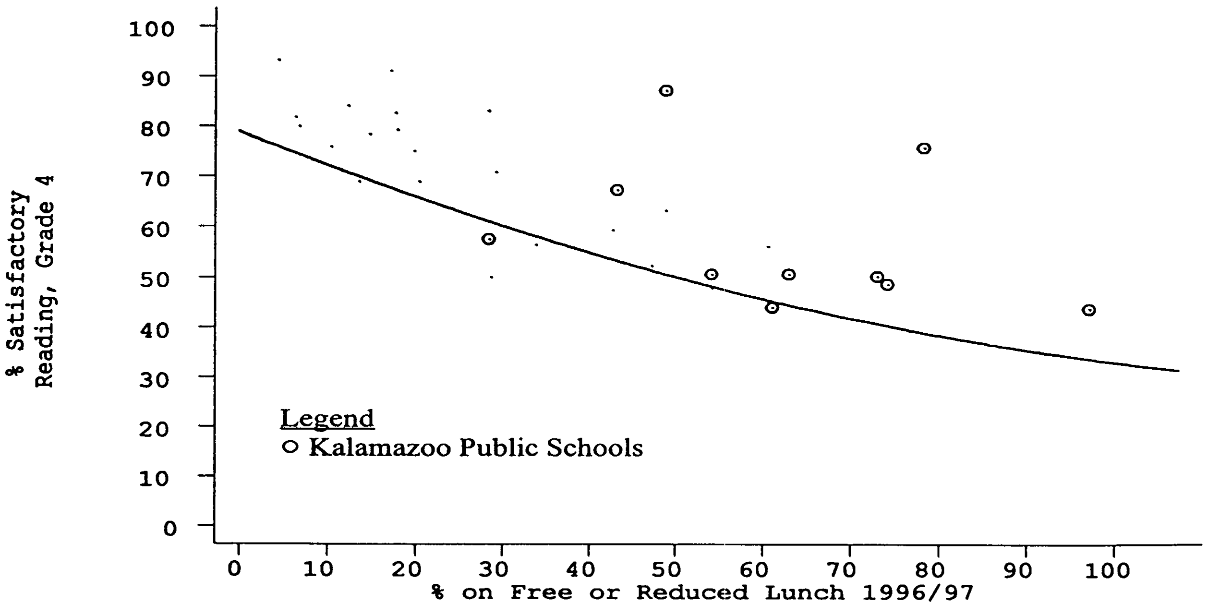

Figure 4. Kalamazoo County 4th Grade Reading Results Using Most Recent Year's Data

case, the y-intercept is 78.91 percent. The sensitivity of the typical school's passing rate on this test to its rate of student eligibility for free- or reduced-price lunch decreases as the poverty rate increases. The decreases in overall passing rates that occur with a one percentage point increase in poverty for typical schools that have free- or reduced-price eligibility percentages of 10 percent, 50 percent, and 90 percent are .650, .461, and .271 percentage points, respectively. The statistical analysis indicates that as buildings have larger and larger percentages of students in poverty, there is less and less of an impact on passing rates.

Figure 4 shows that two KPS elementary schools, one that has a free- or reduced-price lunch eligibility percentage of about 50 percent and one that has a percentage of about 80 percent, exceeded the norm in 1997/98, and had statistically significant positive residuals. The scatter plots in both figures 3 and 4 again show considerable variability among buildings.

One aspect of the shapes of the regression curves in figures 1 and 3 is that they are concave upward. (In figure 1, this is true up to a free- or reduced-price lunch percentage of about 70 percent.) This means that a line segment connecting any two points along the curve lies entirely above the curve. A result of this shape is that if you conduct the "thought experiment" of combining two "typical" schools lying on the regression line, the predicted passing rate would be lower than the sum of the two schools. In other words, consider two typical schools with free- or reduced-price eligibility percentages of 10 percent and 50 percent. The predicted passing rate of a school that averages these two free-or reduced-price lunch eligibility rates–i.e., has a 30 percent eligibility rate—is lower than the average passing rates of the two schools.[8]

---

[8]The parameter estimates in the top panel of table 1 indicate that the predicted passing rates of schools with 10 percent and 50 percent free- or reduced-price lunch percentages are 63.33 and 41.74 for the reading test. The average of these two rates is 52.54. The predicted passing rate of a school with a free- or reduced-price lunch percentage of 30 percent is 49.67.

Math. Table 1 shows that the best model for math passing rates is linear for both the four-year average passing rate and the 1997-98 passing rate. Figures 5 and 6 provide the regression line and the scatter diagram for all buildings in the sample and for buildings only in Kalamazoo County when the four-year average passing rate is the dependent variable. The estimated constant suggests that the y-intercept is 81.68 percent; which, in other words, is the predicted passing rate of the typical school if it were to have no students eligible for free- or reduced-price lunches. The estimated slope coefficient of -.48 indicates that the sensitivity of the math MEAP's passing rate to the percentage of disadvantaged students does not depend on the level of the free- or reduced-price lunch percentage. Anywhere on the regression line, a one percentage point increase in poverty leads to a predicted decrease in the passing rate of .48 percentage points.

The scatter diagram for the math passing rates is just as variable as the scatter diagram for reading results. In figure 5, it is interesting to note that there are three schools with free- or reduced-price lunch percentages of over 90 percent that have four year-average passing rates in math of 70 percent or higher.

The lower panel of table 1 and figures 7 and 8 show the results when the 1997-98 test results are used. The constant term is greater–88.3 percent compared to 81.68 percent—meaning that the predicted passing rate for the typical school if it were to have no disadvantaged students is higher than in the model estimated with averaged data. The slope coefficient of -.40 is smaller (in absolute value) than in the upper panel, meaning that the typical school's passing rate is less sensitive to its free- or reduced-price lunch percentage. A one percentage point increase in poverty reduces the typical building's passing rate in 1997-98 by .40 percentage points.

Figure 8 shows that, again, two KPS schools performed above expectations in 1997/98.

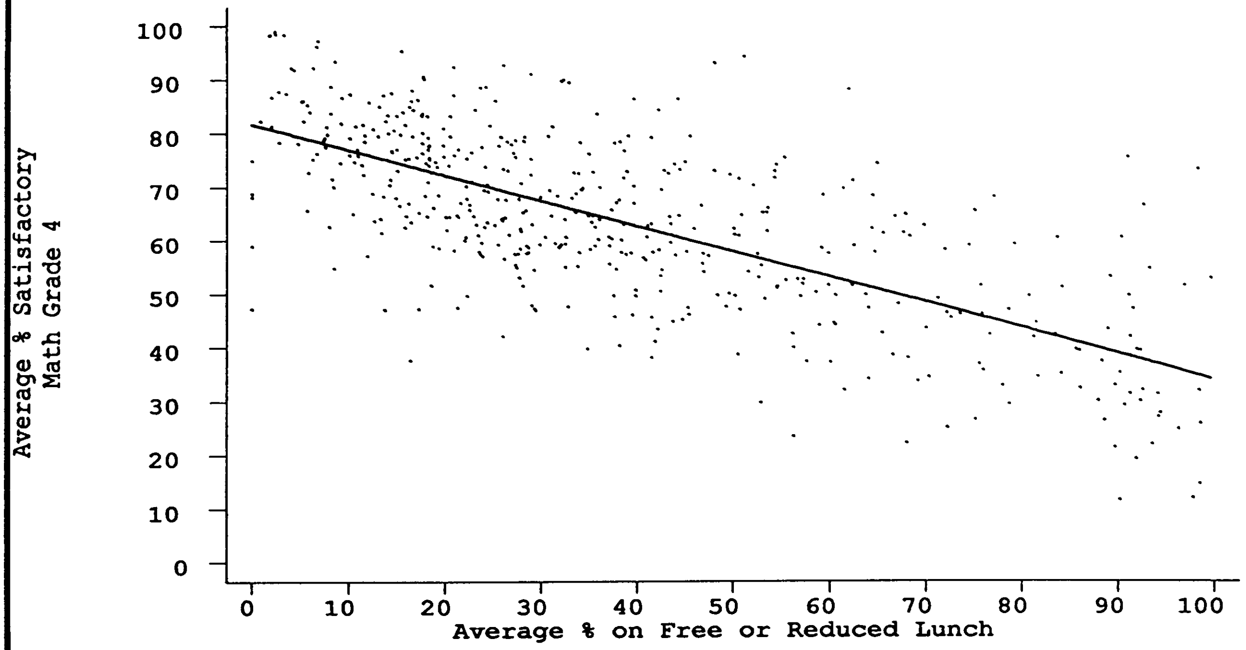

Figure 5. 4th Grade Math Results Using Averaged Data

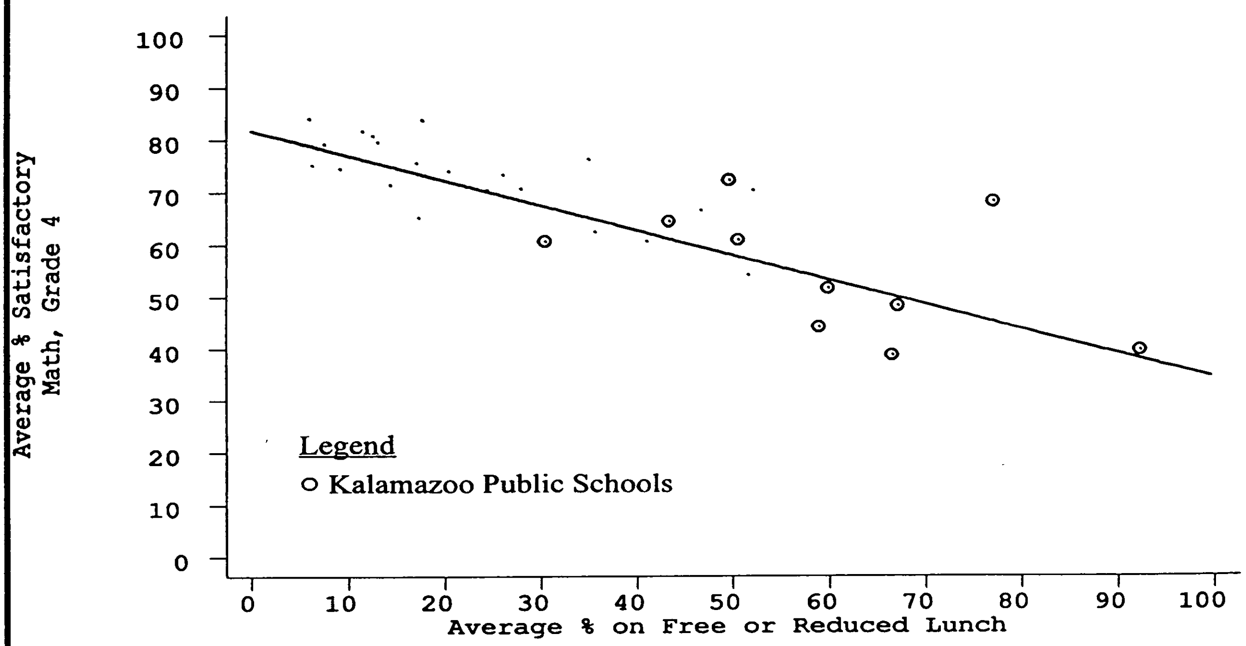

Figure 6. Kalamazoo County 4th Grade Math Results Using Averaged Data

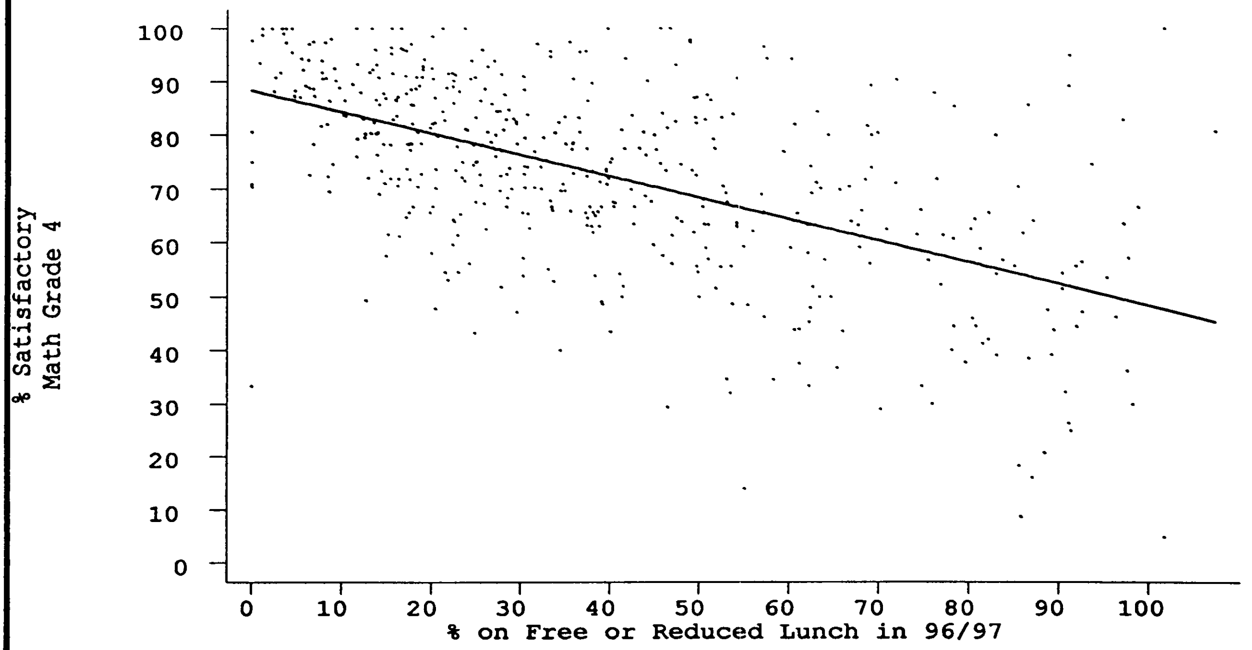

Figure 7. 4th Grade Math Results Using Most Recent Year's Data

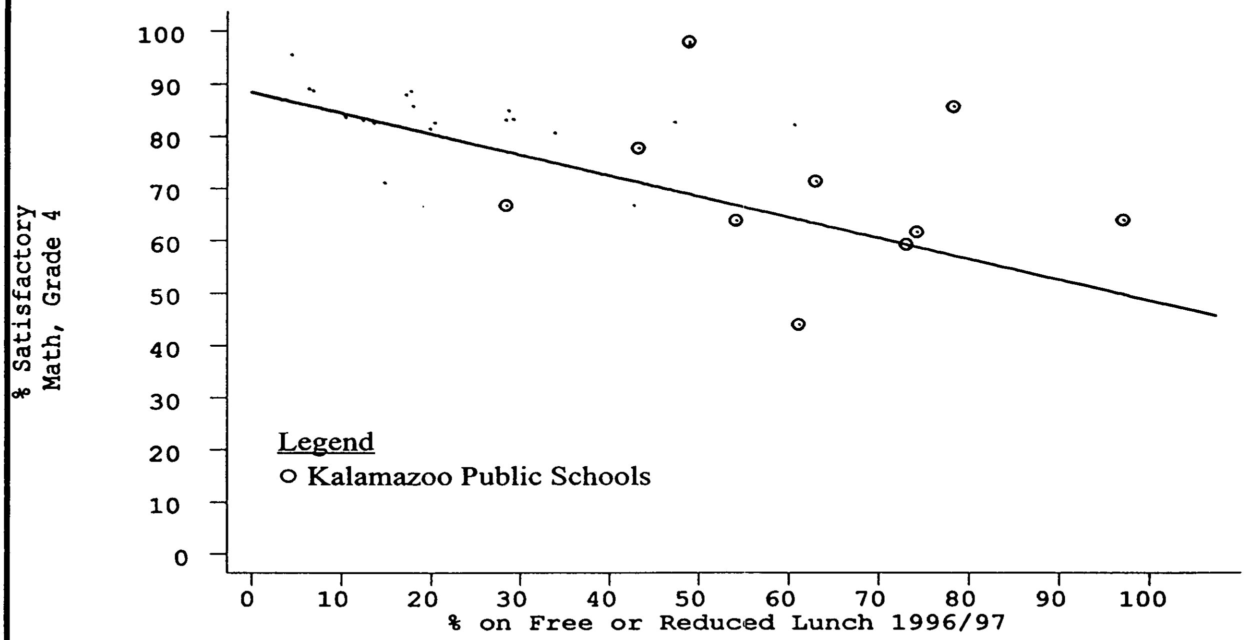

Figure 8. Kalamazoo County 4th Grade Math Results Using Most Recent Year's Data

Science. The results of the regression analyses for (5^{th} grade) science MEAP passing rates are similar to the results for math, except that the passing rates are much lower. Table 1 demonstrates that, as with math, the best functional form for science was linear for both the four-year average data and for the 1997/98 data. The top panel of table 1 and the graphs in figures 9 and 10 show that the y-intercept, the predicted passing rate for the typical school if it were to have no poor students, was 59.65 percent. The estimated slope coefficient of -.44 implies that the predicted passing rate for a school building decreased by .44 percentage points for every one percentage point increase in the free- or reduced-price lunch eligibility percentage.

The scatter diagram in figure 9 appears to show less variability around the regression line for the (four-year average) passing rates on the science test than there was for reading or math. Only a few buildings seem to be a significant distance away from the line. Indeed, figure 10 shows that the buildings in Kalamazoo County are all very close to the predicted line.

The bottom panel of table 1 and the graphs in figures 11 and 12 indicate that the y-intercept for the regression analysis using 1997-98 data was 56.19 percent and that the slope coefficient was -.49. In other words, the typical school that had no students eligible for free- or reduced-price lunches in 1996-97 had a predicted passing rate of about 56 percent on this test, and the typical building's passing rate declined by about .49 percentage points for each percentage point increase in poverty rate.

Interestingly, the variability in figure 11 appears to be much greater than in any of the previous scatter diagrams suggesting that the buildings' results on this test were all over the map. Apparently some schools experienced very large improvements in the science test results in 1997-98. In figure 12, all of the KPS schools were close to the regression line, i.e., performed about as well

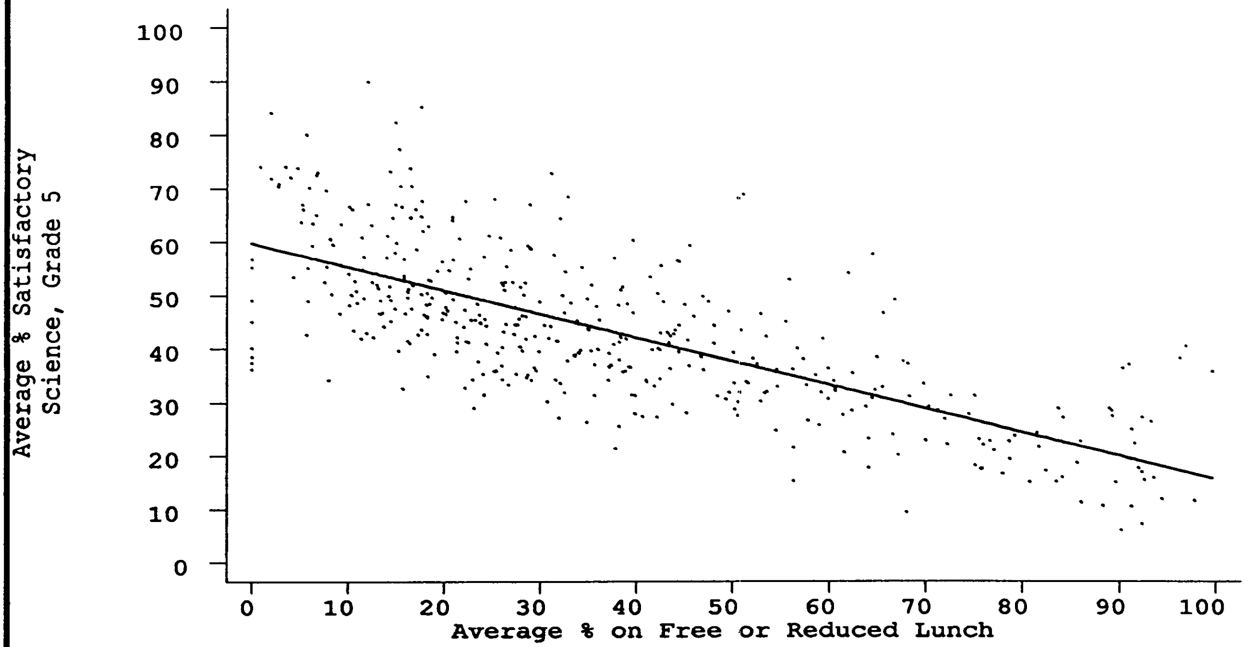

Figure 9. 5th Grade Science Results Using Averaged Data

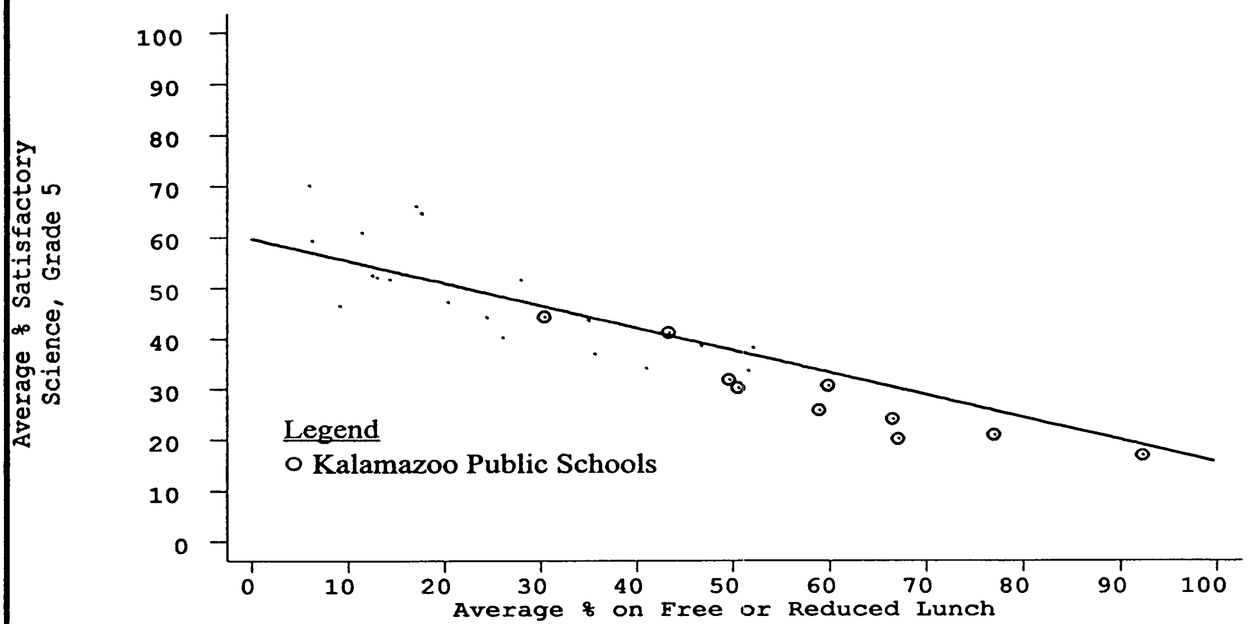

Figure 10. Kalamazoo County 5th Grade Science Results Using Averaged Data

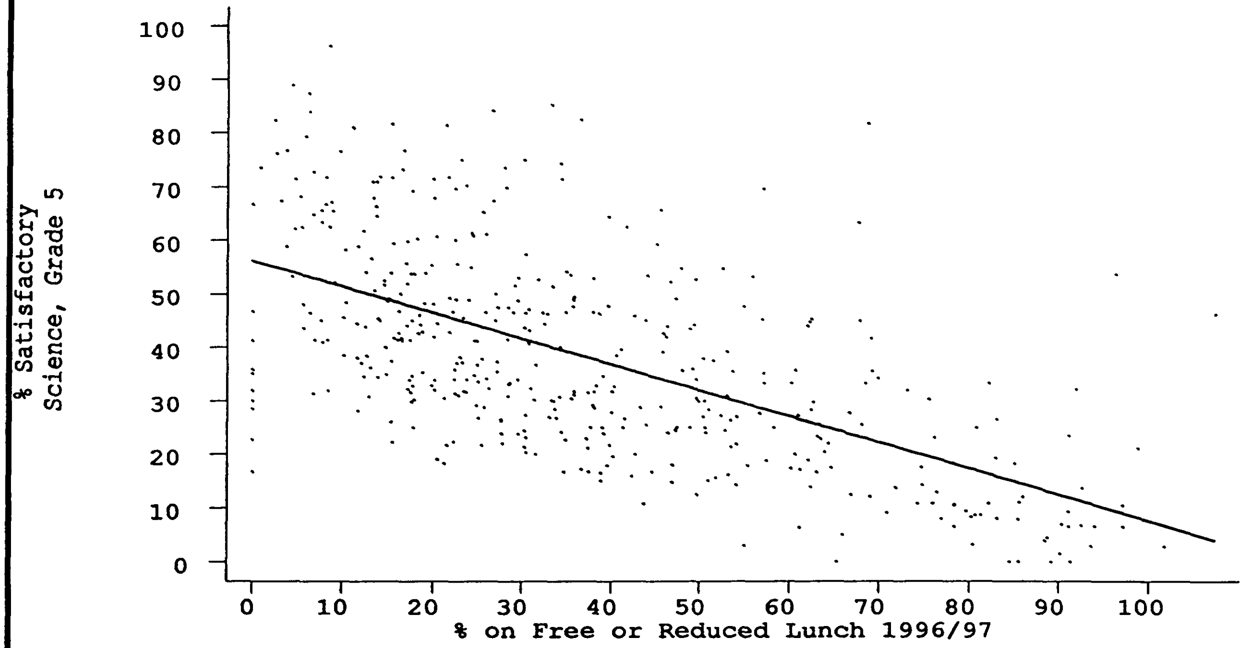

Figure 11. 5th Grade Science Results Using Most Recent Year's Data

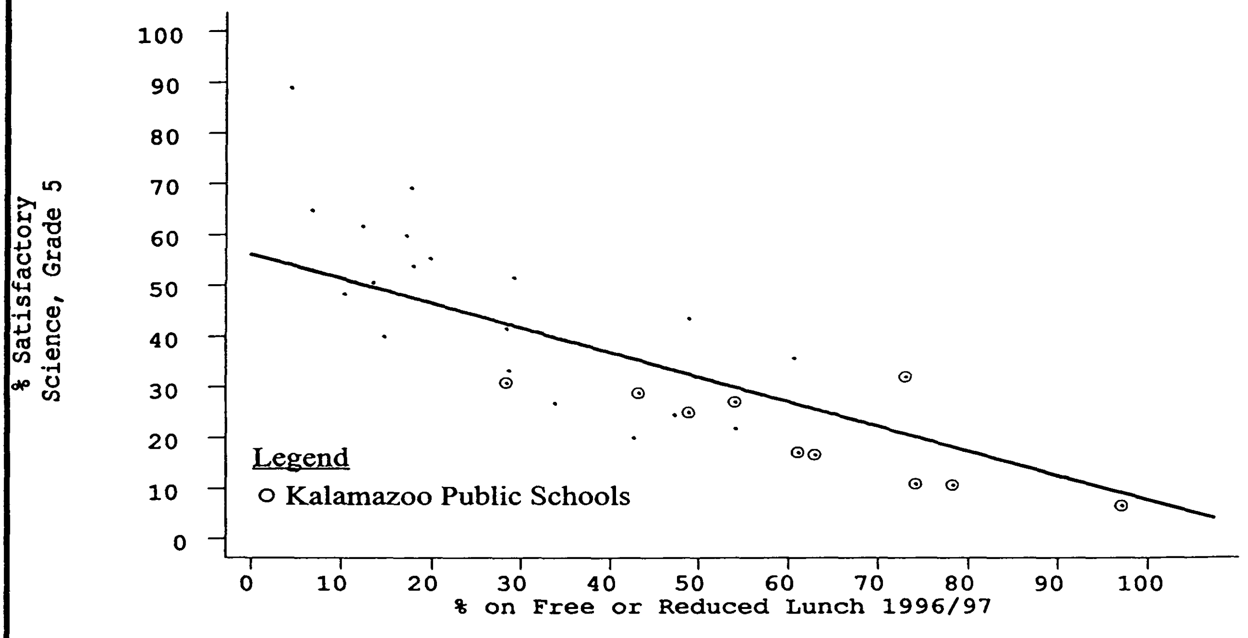

Figure 12. Kalamazoo County 5th Grade Science Results Using Most Recent Year's Data

as expected given their free- or reduced-price lunch percentages. However, there was one building in the county outside of KPS that was significantly above the line (a Portage school).

Writing. The 5th grade writing results were similar in functional form to the 4th grade reading tests. The top panel of table 1 shows that a cubic equation best fit the three-year average data, and a quadratic best fit the 1997/98 data. Figures 13 through 16 plot the results. The y-intercept and the slope coefficients for the three-year average data at free- or reduced-price lunch percentages of 10 percent, 50 percent, and 90 percent are 83.82 percent (y-intercept) and -.775, -.336, and -.638 (slope coefficients), respectively. This means that the three-year average passing rate for the typical school with no poor students was 83.82 percent, and the predicted sensitivity of the average passing rate to the percentage of students who were eligible for free- or reduced-price lunches was quite high at low levels of poverty and at high levels of poverty, but was much lower in between for the average building. The decrease in the predicted average passing rate on the writing test for a one percentage point increase in the free- or reduced-price lunch eligibility was .775 percentage points for schools with around 10 percent eligibility; .336 percentage points for schools that have 50 percent eligibility; and .638 percentage points for schools with an eligibility rate of 90 percent. Figure 14 shows that all of the county school buildings were near the regression line.

The regression estimates using the latest year's passing rates imply a y-intercept of 81.41 percent, which would be the predicted passing rate for a building with no students who were eligible for free- or reduced-price lunches. The slope coefficients work out to be -.570, -.373, and -.176 at free- or reduced-price lunch eligibility rates of 10 percent, 50 percent, and 90 percent, respectively. In other words, the sensitivity of the pass rates to having more students eligible for free- or reduced-price lunches decreases as the percentage increases. For schools with 10 percent poverty, the decline in 1997-98 passing rates was .57 percentage points for a one percentage point increase in free- or

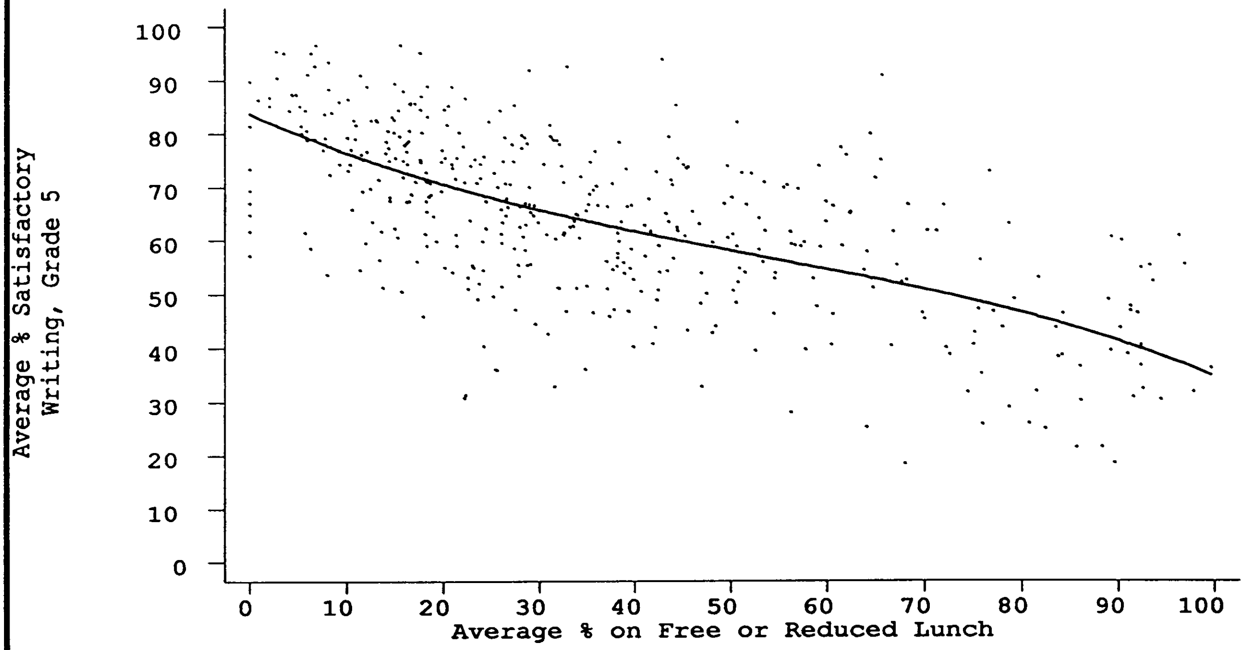

Figure 13. 5th Grade Writing Results Using Averaged Data

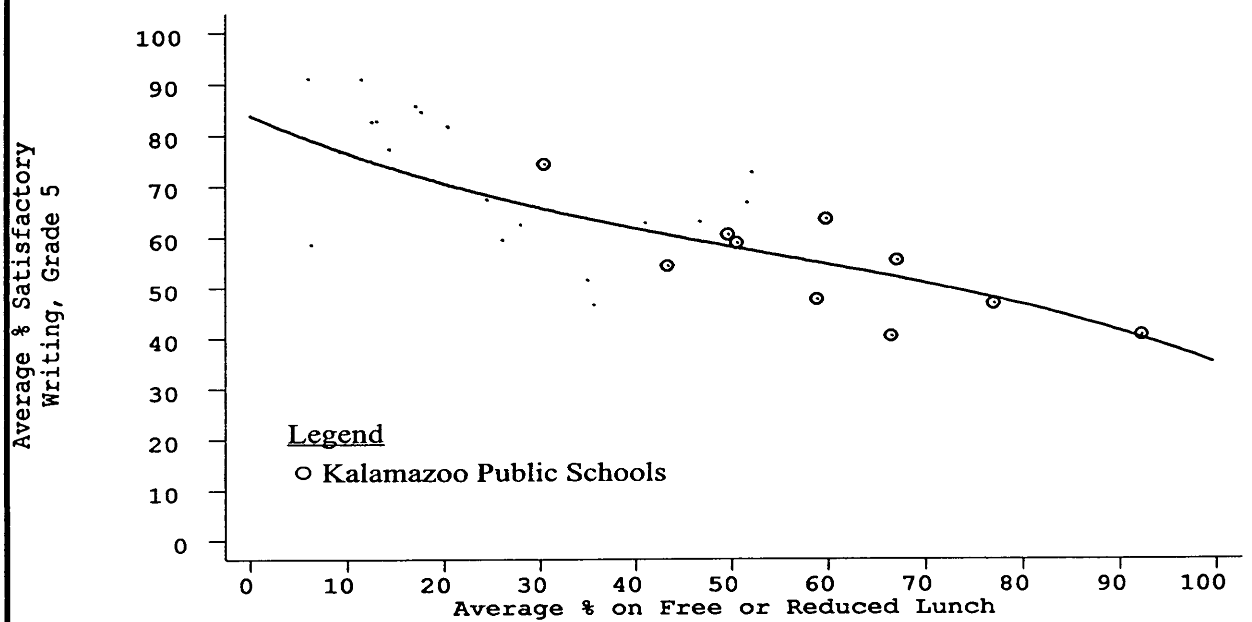

Figure 14. Kalamazoo County 5th Grade Writing Results Using Averaged Data

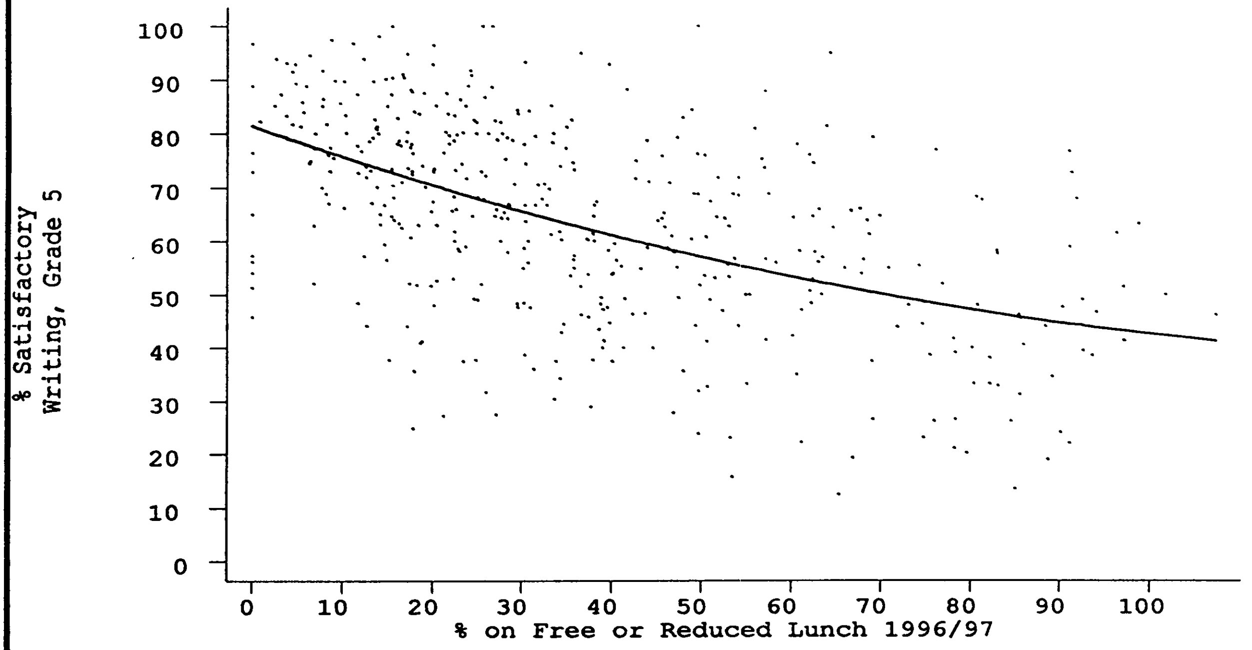

Figure 15. 5th Grade Writing Results Using Most Recent Year's Data

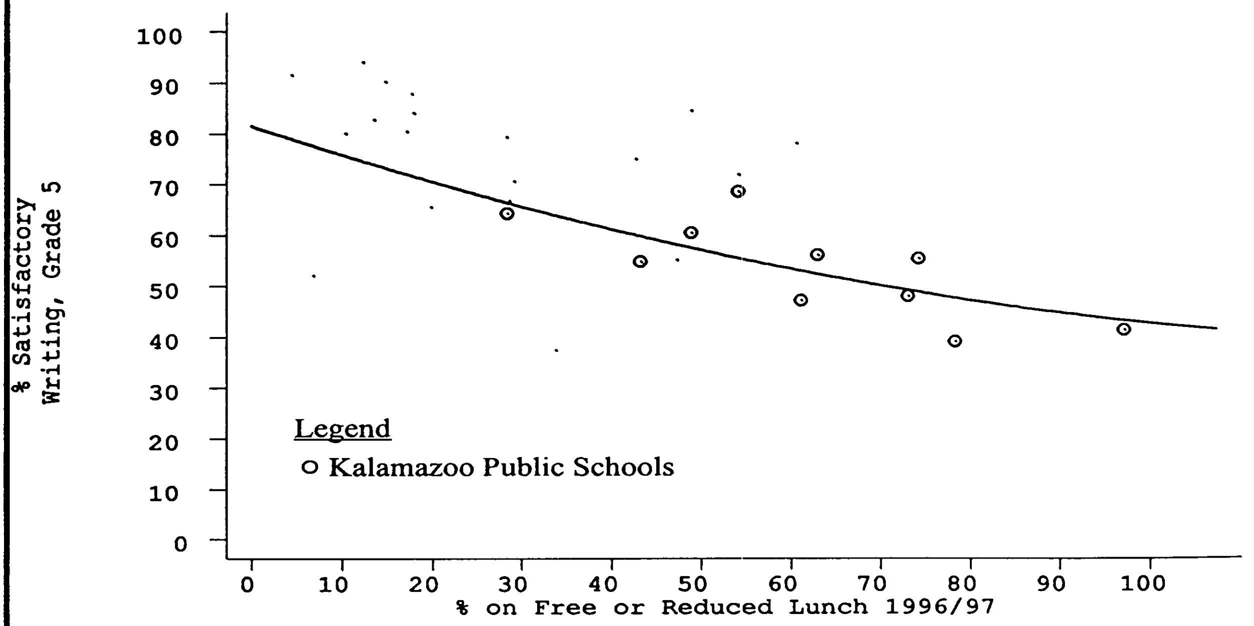

Figure 16. Kalamazoo County 5th Grade Writing Results Using Most Recent Year's Data

reduced-price lunch eligibility. It was .373 percentage points at a school with 50 percent poverty, and it was .176 for the typical school that had 90 percent of its students who were poor.

Figure 16 shows that all of the KPS schools scored near the regression line for the writing test in 1997-98. However in the county, two schools (both in Comstock) did significantly better than predicted and two schools did significantly worse.

In summary, the statistical analyses of these test score data confirm a negative relationship between how well a building does on the MEAP and its free- or reduced-price lunch eligibility percentage. But in looking at the scatter diagrams, it is always the case that there are schools with relatively high levels of poverty that perform as well as or better than schools with much lower levels of poverty. Similarly, there are schools with very low levels of poverty that perform much worse than buildings with much higher levels of poverty. In the absence of other information, one can predict that higher levels of student eligibility for free- or reduced-price lunches at a school will increase the likelihood that test performance is lower. However, it is far from being a perfect predictor, and individuals comparing schools would be advised to (1) look at other data besides the free- or reduced-price lunch eligibility percentage and (2) find out where the schools' test results are relative to where they would be predicted.