The Effects of School Location on Math and Science Achievement Trends: A Primer on Growth Modeling in Education Policy Research

Occasional Research Paper, No. 11

Edward B. Reeves

Morehead State University

Edward B. Reeves is Professor of Sociology and Director of the Center for Educational Research and Leadership, Institute for Regional Analysis and Public Policy. An earlier version of this paper was presented at the Twelfth International Conference on Learning, convened in July 2005 at the University of Granada in Granada, Spain. Direct inquiries to the author at Institute for Regional Analysis and Public Policy, Morehead State University, Morehead, KY 40351, or via e-mail to: e.reeves@moreheadstate.edu.

1

The Effects of School Location on Math and Science Achievement Trends: A Primer on Growth Modeling in Education Policy Research

Abstract

The effect of school location on math and science learning is currently an important policy issue in the United States and in other countries, such as Australia. The present paper uses a 5-year series of math and science achievement data from the state of Kentucky to determine the effects of school location on learning in these subject areas. Adopting an organizational assessment approach, I show how growth models may be used to estimate achievement trends. I also demonstrate methods for discovering two important sources of invalidity in growth models: regression artifacts and spuriousness. Failure to account for these sources of invalidity may lead to erroneous policy conclusions. Two examples of growth models are provided—a linear model and a nonlinear model. The results of these analyses do not support the common contention that there is a rural achievement gap in math and science. One implication of these findings is that, if policymakers wish to enhance math learning, they will accomplish this more effectively by interventions and programs that increase the motivation and opportunity to learn among low-income students, regardless of school location. Because current U.S. education policy is focused on documenting "adequate yearly progress" in schools, growth modeling is likely to become the preferred methodology of policy researchers.

The Effects of School Location on Math and Science Achievement Trends: A Primer on Growth Modeling in Education Policy Research

During the past 20 years, education policy research in the USA has focused on tracking student and school achievement trends for the purposes of insuring accountability, making equity adjustments in school resources, and ascertaining the effectiveness of learning interventions (Ladd, 1996; National Research Council, 2001). This direction in education policy research was inaugurated by education reform initiatives in states like Minnesota, Texas, and Kentucky and was subsequently codified as national policy by the U.S. No Child Left Behind Act of 2001.

This paper examines growth modeling, an important research tool that is available to researchers who want to inform policymakers about school achievement trends. Growth modeling is adaptable to the investigation of a broad array of policy-related issues, and it is equally useful for experimental and quasi-experimental research designs (Nye, Konstantopoulos, & Hedges, 2004; Raudenbush & Bryk, 2002). Although growth models may be used to study student-level achievement, for example, the relationships between where students start and how rapidly they progress (Seltzer, Choi, & Thum, 2003), in this paper I adopt an organizational assessment perspective that focuses on grade-level achievement trends (Reeves & Bylund, 2005). I will demonstrate how growth curves of math and science achievement may be estimated. In addition, I will show how to test for: (1) the effects of a policy-relevant covariate—school location—on the growth curve parameters, (2) the effect of a regression artifact (regression toward the mean), and (3) spuriousness of the results. Since learning trends are not always monotonic, I will offer an example of a nonlinear (quadratic) growth model following my example of a linear growth model.

The empirical issue dealt with in the examples presented below is an important one for education policy. It concerns the effects of school location on math and science achievement trends. A school cannot usually change its location, yet location conceivably may have consequences for how well students learn at the school. The investigation of a rural gap in academic achievement has been recently explored in studies conducted in the USA (Fan & Chen, 1999; Greenberg & Teixeira, 1995; Roscigno & Crowley, 2001; Reeves & Bylund, 2005), in Australia (Webster & Fisher, 2000; Young, 1998, 2000), and in cross-national comparisons (Williams, 2004). In the USA, rural education is associated with disadvantage in the public discourse. While research on this matter has not yielded consistent results (Fan & Chen, 1999; Roscigno & Crowley, 2001; Israel, Beaulieu, & Hartless, 2001), it is reasonable to hypothesize that, if rural disadvantage does exist, it is likely to be found in significant learning gaps in mathematics and science. Rural schools are disproportionately likely to have an inadequate pool of teachers qualified in these subjects and insufficient funds to maintain up-to-date computers, instructional software, and laboratory facilities (Education Alliance, 2004; Williams, 2005).

I will examine a 5-year time series of math and science achievement trends in Kentucky public schools to find out if a significant rural school achievement gap exists. If the growth models indicate such a gap, I will then determine (a) how much of the gap may be attributable to a regression artifact and (b) if the gap is spurious (that is, it can be better explained by a factor other than school location). Limited space does not allow for

an exhaustive analysis of all Kentucky grade levels that are tested each year in math and science. Therefore, I have chosen examples of two growth models that have instructive features. These analyses, shown below, were estimated with Hierarchical Linear Modeling software (HLM 6).[1]

Eyeballing Achievement Trends using Graphs

The first step in gaining an understanding of achievement trends is to plot them. The growth curves plotted in Figure 1 show that, from 1999 to 2003, 5th grade math achievement increased consistently in rural and nonrural Kentucky public schools alike. These plotted trends also indicate the presence of approximately a 7-point gap between the rural and nonrural schools at the start of the series. By 2003, this gap had closed to around 4 points. An examination of these plots may lead us to ask questions like the following: Is the initial rural gap large enough to be statistically significant? How significant is the apparent closing of the gap in later years?

The growth curves plotted in Figure 2 of 11th grade science achievement tell a somewhat different story. In 1999, the rural school gap in science achievement appears to be slightly greater than 2.5 points; and, although there are fluctuations in the trend lines, the gap appears little changed in 2003. In these examples, however, in contrast with 5th grade math scale scores where the trends are monotonic, the 11th grade science scale scores appear to level off in rural schools and are actually on a downward trend in nonrural schools by the end of the series. The plotted trends raise additional questions: Is the rural school gap in 11th grade science achievement statistically significant? Is the slowdown or reversal in the rate of growth after 2002 significant? Is the decline in science achievement indicated for nonrural schools in 2003 significantly greater than the slowdown indicated for rural schools, which occurs at the same time?

Eyeballing growth curves in this manner leads us to ask questions like the ones posed above. Growth modeling has the great advantage that it will provide precise answers to these questions. There are other questions that must be answered if policymakers are to make informed decisions about what the trends in achievement really mean. These questions include: What is the effect of regression toward the mean on the trends depicted in Figures 1 and 2? Are the apparent rural/nonrural differences shown in the graphs the result of spuriousness? Growth modeling, again, will provide precise answers to questions such as these.

---

[1] The 5th grade math and 11th grade science scale scores that are used in this study were acquired from the Kentucky Department of Education. KDE also provided data on the annual percentage of students taking the tests who were eligible for free or reduced-price lunch. Information about Kentucky's testing program may be obtained by visiting the following web site: http://www.edication.ky.gov/KDE. The measurement of school location is a dummy variable (rural = 1; nonrural = 0). "Rural location" means that the school is in a community with less than 2,500 inhabitants. This variable was extracted from the Common Core of Data, National Center for Educational Statistics, U.S. Department of Education, and matched to Kentucky public schools.
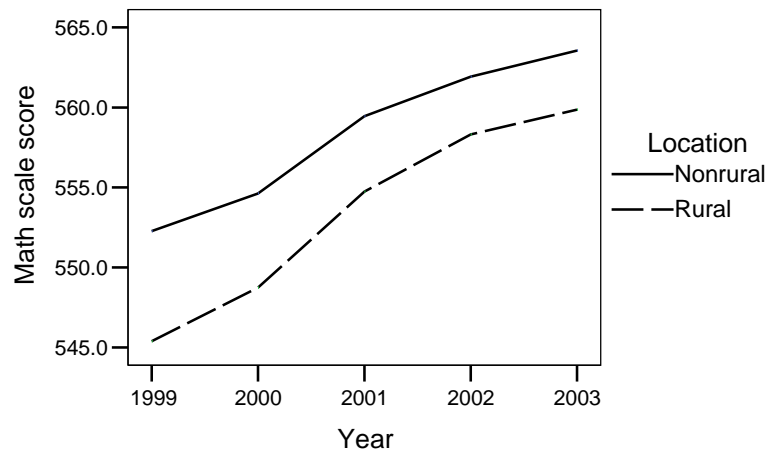
Figure 1
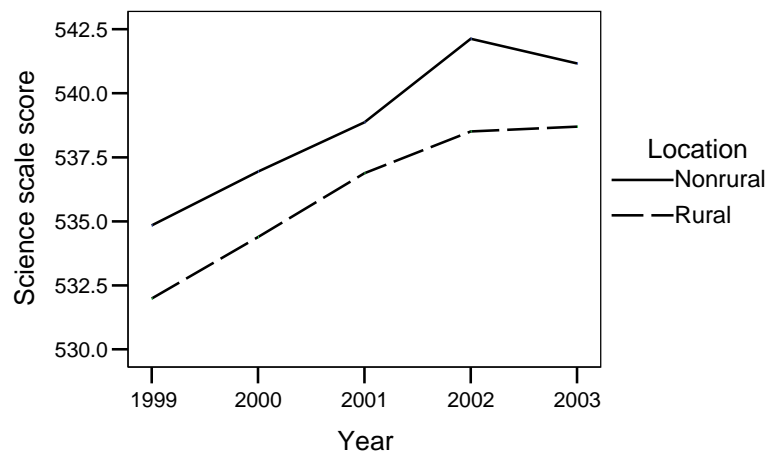Kentucky 5th Grade Math Achievement Trends

Figure 2
Kentucky 11th Grade Science Achievement Trends

An Example of a Linear Growth Model

A second look at the trend lines plotted in Figure 1 will illuminate the parameters that a linear growth model must estimate. First, it must estimate the starting point of the trend. This is the mean scale score for 5th grade math in 1999. I refer to this parameter of the growth model as 'achievement status'. Second, the model must estimate the annual average rate of growth in the math score throughout the time series—in other words, the mean annual increase in the scale score from 1999 to 2003. I term this second

parameter 'achievement growth'. If the plot of the time series of data indicates an approximately linear (i.e., straight-line) rate of growth, then no other parameters need to be estimated in order for the growth model to describe the trend. Although the trend lines shown in Figure 1 are not absolutely straight, they always move upwards and do not reverse direction or display a marked tendency to level off. Thus, a linear growth model will be adequate to estimate such growth curves. (If we suspected otherwise, there is a means to test for nonlinearity of trends as I will show in the second example.)

I will not provide a lengthy explanation of how growth models are estimated with HLM. Interested readers will want to consult several excellent treatments of this topic (Luke, 2004; Raudenbush & Bryk, 2002; Singer & Wilbert, 2003). My purpose here will be to present the results of the HLM growth model analysis, discussing the estimated parameters and what they tell us about the policy-relevant issue of a rural achievement gap in math and science learning in Kentucky public schools.

Table 1A presents the descriptive statistics for the variables that I use to analyze the growth curves for 5th grade math achievement. The three variables shown in the upper panel of Table 1A are within-school, time-dependent variables. This means that they take on different values for each year in the time series. The variable listed first is the school's math scale score for 5th graders. In Kentucky, the math scale scores are constructed following Item Response Theory methods and have a theoretical variation between 325 and 800 points. As Table 1A shows, the actual range in scale scores across the 5-year time series is from 488 to 620 points, with a mean of 556.1 and standard deviation of 17.6. The next variable shown in the top panel of Table 1A is the year index. This variable is coded in integer values that range from 0 to 4, marking the position in the time series. Therefore, the starting year, 1999, is coded 0; the final year, 2003, is coded 4. Years are coded in this manner so that the HLM-generated models will estimate 'achievement status' (the intercept) for the starting year and 'achievement growth' (the year slope) for the annual gain in math scores.

Percent low-income students is the final time-dependent variable shown in the upper panel of Table 1A. This variable refers to the percent of students taking the 5th grade math test in a given year who are eligible for free or reduced-price lunch. Percent eligibility for participation in the federally subsidized lunch program is frequently used in the USA to indicate the proportion of low-income students. From 1999 to 2003, the average of such students at the 5th-grade level in Kentucky public schools was 52.3 percent, with standard deviation of 22.2. Across all schools in the sample, the range of the proportion of low-income 5th graders varied in the extreme, from zero to 100 percent.

The lower panel in Table 1A contains a single time-invariant, between-school variable. This variable is binary, coded 1 if the school is in a rural location and 0 if not. For the purposes of this study, school location is considered time-invariant because it is very unlikely that a school's classification by location has altered during the 5-year period for which I have data to analyze. Of course, the analysis of trends across longer time intervals of (say) 10 to 20 years could very well require that school location be treated as a time-dependent variable. As Table 1A shows, there are 661 Kentucky public elementary schools for which there is time-series data on 5th grade math scale scores, and 44 percent of these schools are rural.

I construct three linear growth models to estimate the effects of the variables discussed above. The first of these models estimates only the basic growth curve

Table 1A

Descriptive Statistics for 5th Grade Variables

Withnin-school variables (N = 3284)	Mean	SD	Min.	Max.
Math scale score	556.1	17.6	488	620
Year index	2	1.4	0	4
% low-income students	52.3	22.2	0	100
Between-school variables (N = 661)				
Rural location	0.44	0.50	0	1

Table 1B

Linear Growth Models for 5th Grade Math Achievement Trends

Parameters	Model 1	Model 2	Model 3
Achievement status (intercept)	549.4**	552.4**	550.6**
Rural location	--	–6.7**	–1.8
Achievement growth (year slope)	3.4**	3.0**	2.9**
Rural location	--	0.9**	0.5
% low-income students (avg. effect)	--	--	–0.3**
% low-income students x year	--	--	0.0
Goodness of fit X^2(df)	870.4(2)**	899.6(2)**	1223.1(2)**

*p<.05 **p<.01

parameters of achievement status and achievement growth. The second model estimates the growth curves depicted in Figure 1 and tests the effects of rural location on achievement status and on achievement growth. The third model tests for spuriousness of the results obtained from the second model.

Turning to Model 1, which is shown in Table 1B, we find that achievement status (the mean 5th grade math scale score in 1999) is estimated to be 549.4, while achievement growth is estimated to average 3.4 per annum. From these results we can estimate, by extrapolation, that the mean 5th grade math score in 2003 is 563, 13.6 points above the achievement status benchmark. Not only is this rate of growth in math achievement statistically significant, the effect size is substantively important. The annual growth rate is equal to 0.2 standard deviation (3.4/17.6 = 0.2).

Model 2 tests the effects of rural location on achievement status and achievement growth, and therefore estimates the trends displayed in Figure 1. In 1999, rural 5th graders are found to score an average of 6.7 points below comparable students in nonrural schools. The average math score of the nonrural students is 552.4. With respect to achievement growth, Model 2 indicates that the rural 5th graders are gaining at a rate that is statistically significantly greater than the nonrural 5th graders. The achievement growth in nonrural schools is estimated to be 3.0, while in rural schools it is 3.9 (3.0 + 0.9). The results in Model 2 might be used to support the conclusion that, although 5th graders in rural schools started out behind their nonrural counterparts, this gap is closing by 13 percent (0.9/6.7 = 0.13) each year. If this rate of improvement continues, it will take rural 5th graders only eight years to catch up with and even to surpass their nonrural peers in math achievement. However, this optimistic projection may be erroneous and could lead to incorrect policy decisions if taken at face value. It is

possible that the results obtained in Model 2 artificially inflate the rural growth rate because of a commonly overlooked regression artifact.

Regression toward the mean poses a problem whenever repeated measurements are taken on a sample, as is the case here where 5th grade math achievement is being measured annually. 'Regression toward the mean' refers to the tendency for cases in the sample that are in the tails of the distribution when first measured to be found closer to the mean when measured later in time. Campbell and Kenny (1999) offer numerous examples of this regression artifact and how policy research can be seriously misled by a failure to take it into account. In the present study, the effect of the regression artifact leads to the result that schools which scored either above or below the mean math scale score in 1999 are likely to score closer to the mean in a subsequent year. As a consequence, schools scoring lower than the mean in 1999 will appear to improve at an above-average rate in later years, while schools scoring higher than the mean in 1999 will appear to improve at a below-average rate. Because rural schools have been shown to score significantly below nonrural schools in 1999, their positive achievement growth is confounded with this effect of regression toward the mean. Growth analysis with HLM gives us a means to assess how much the regression artifact may affect the results obtained by Model 2.

A thorough explanation of how to test for the regression artifact with HLM is beyond the scope of this paper. Interested readers will want to consult Raudenbush and Bryk (2002, pp. 361-364; see also, Seltzer, Choi, & Thum, 2003). Following the procedures that they describe, I add a parameter to the HLM equation that represents the effect of achievement status on achievement growth (not shown), which simulates regression toward the mean. We may then determine if controlling this effect influences the estimated effect of rural location on the achievement growth coefficient. The result that I obtained is that the rural location coefficient is reduced by a statistically significant amount, from 0.9 (shown in Model 2) to 0.5, a reduction of 44 percent. Although this adjusted rural location effect remains statistically significant at the 0.05 level, it compels us to revise the estimation of how rapidly the rural 5th graders are improving their math scores relative to the nonrural students. The revised achievement growth for rural 5th graders is 3.5, and at this rate it will take them about 13 years to catch up, instead of 8 years as previously forecast. In summary, by testing for the effect of regression toward the mean, it was shown that nearly half of the apparent growth advantage of rural location, which is presented in Model 2, is not real but is the result of a regression artifact.

Model 3 in Table 1B tests another possible source of invalidity for the results obtained by Model 2. I refer to the possibility that the significant effects of rural location on achievement status and achievement growth may be spurious. That is, the relationships that were found to be significant in Model 2 could be the result of an omitted variable, one that is not included in the model. Research in Kentucky (Reeves, 2000, 2003) and elsewhere (Roscigno & Crowley, 2001; Williams, 2005) finds academic achievement is often quite sensitive to students' family incomes. Furthermore, families of rural students earn on average lower incomes than the families of nonrural students. The association between rural location, low family income, and low academic achievement suggests the hypothesis that the relationships found between rural location and achievement status and achievement growth (Model 2) may be spurious.[2] To test this possibility, two within-

---

[2] Spuriousness is a separate issue from regression toward the mean and must be tested separately. In this instance, spuriousness is an issue regardless if we accept the results

school parameters are added in Model 3: the first parameter represents the average overall effect of the percent of low-income students on the mean math scale score; the second parameter is an interaction that represents the effect of low-income students on the annual change in the score.

The results after adding these variables are shown in Model 3 (Table 1B). The average low-income student effect and the interaction effect together sharply reduce the size of the rural location coefficients and render them both statistically insignificant. This is strong evidence that the earlier findings of significant rural location effects are indeed spurious. The proportion of low-income students, rather than rural location, is the significant factor that affects 5th grade math achievement trends. In absolute terms, the effect size of low-income students is about twice the effect size of rural location: One standard deviation increase in the percent of low-income students is associated with 0.4 standard deviation decrease in the average math score. The effect of low-income students on the annual change in the math score, as revealed by the interaction, is nil, however. Therefore, the effect of low-income students is constant across the time series. This being the case, regression toward the mean may not be an issue with respect to the effect of the low-income student variable.[3]

An Example of a Nonlinear (Quadratic) Growth Model

The previous example of growth modeling estimated trends that were deemed linear. The next example deals with trends that are curvilinear and, therefore, require the addition of a squared (quadratic) term to the growth modeling equation in order to account for this nonlinearity. Only two models will be estimated in this example. The reason for this will become clear. Table 2A displays the descriptive statistics for the variables used in this example. Two within-school, time-dependent variables are shown. These are 11th grade science scale scores and the year index. Descriptive statistics for the quadratic term, which is simply the square of the year index, are not shown in the table. The mean science scale score for this sample is 537.8 with standard deviation of 12.2. Once again, the theoretical range of these scores as determined by IRT methods varies from 325 to 800. The actual range in the scores is from 481 to 577. The year index is calculated as in the previous example. Rural location is the only between-school variable, and, as before, it is assumed to be constant across the 5-year time series. Among the 217 high schools whose 11th graders were tested for science achievement, 34 percent were located in rural areas.

Table 2B presents the results of estimating nonlinear (quadratic) growth models using the variables just described. Model 1 shows three growth curve parameters. Achievement status is the same as in the linear growth model example. The other two

---

of Model 2 shown in Table 1B, or the revised results that accounted for the regression artifact.

[3] Technically, there is a regression artifact with low-income students too. Due to regression toward the mean, the effect of the percent of low-income students actually has a negative sign; that is, low-income students really exert a small negative influence on achievement growth. This effect is not significant, however. Consequently, I allow the results in Model 3 to stand.

Table 2A

Descriptive Statistics of 11th Grade Variables

Withnin-school variables (N = 1082)	Mean	SD	Min.	Max.
Science scale score	537.8	12.2	481	577
Year index	2	1.4	0	4
Between-school variables (N = 217)				
Rural location	0.34	0.47	0	1

Table 2B

Nonlinear (Quadratic) Growth Models for 11th Grade Science Achievement Trends

Parameters	Model 1	Model 2
Achievement status (intercept)	533.6**	534.5**
Rural location	--	–2.7
Initial achievement growth (year slope)	3.2**	3.1**
Rural location	--	0.1
Deceleration (year^2)	–0.4**	–0.3**
Rural location	--	0.0
Goodness of fit X^2(df)	222.8(2)**	224.9(2)**

*p<.05 **p<.01

parameters are different, since they address the nonlinearity of the trend estimated by the present model. The year slope now estimates 'Initial achievement growth'. This term refers to the estimated annual rate of increase net of the effect of deceleration (i.e., the decline in the growth rate). 'Deceleration', the effect that is obtained by squaring the year index, is used to test the nonlinearity of achievement growth. When interpreting the growth trend estimated with this nonlinear model, both initial achievement growth and deceleration must be considered together, because the actual annual rate of growth in science achievement is calculated as the sum of these two parameter estimates.

The results of Model 1 in Table 2B estimate an achievement status of 533.6 for 11th graders in the average high school. An average initial achievement growth of 3.2 is also estimated. However, the actual growth rate is less than this because of the effect of deceleration. Deceleration averages negative 0.4 per annum. Therefore, the actual growth rate in science achievement is 2.8 (3.2 – 0.4). Even after deceleration is taken into account, the effect size of the actual growth rate is substantial. The annual increase is 0.2 standard deviation (2.8/12.2 = 0.2). Extrapolating from the estimates in Model 1, 11th graders average 544.8 on the science test in 2003, a gain of 11.2 in four years.

Model 2 in Table 2B estimates the nonlinear growth curves shown in Figure 2 and tests for the influence of rural location on these trends. As the results indicate, the effects of rural location are too small to be statistically significant influences on achievement status, initial achievement growth, or deceleration. Furthermore, while Figure 2 appears to indicate a different deceleration for rural and nonrural 11th graders, the results in Model 2 do not support a significant difference in the two

groups of students. Because the location effects are not significant in this example, I do not test for regression toward the mean or for spuriousness.

Policy Implications and Uses

The importance of addressing resource deficiencies in rural schools is part of the current policy agenda in the USA (United States Government Accountability Office, 2004). While the analyses presented above are intended to be instructive, rather than thoroughgoing in their investigation of location effects on math and science achievement trends in Kentucky, the implications for policy, though limited, are nevertheless clear.

Rural location does not significantly influence the achievement trends of 5th grade math or 11th grade science, at least insofar as these trends are accurately measured by Kentucky's official tests. In the case of 5th grade math learning, the apparently significant rural gap is not real. The gap is explained by the presence of a greater percentage of low-income students in rural schools. Moreover, the negative effect of student poverty on math learning transcends location. To increase 5th grade math performance, policymakers could focus on initiatives that enhance the motivation and opportunity to learn of economically disadvantaged students regardless of where their schools are located. With respect to 11th grade science learning, the effects of location on achievement status and the actual growth rate (after the correction for deceleration) are too small and insignificant to warrant immediate attention by policymakers. My second example did not explore the effects of low-income students on science achievement trends in the 11th grade, because the focus of this study was to assess the effects of school location on these trends. This omission should not be construed to suggest that student poverty has little influence on science achievement scores. Only additional analysis can determine if this is true.

The results presented above do not support the claim that rural students achieve less well than their nonrural peers in math and science. This being the case, it cannot be suggested that rural schools necessarily are deprived of qualified teachers in these subjects or that their learning technology and laboratory facilities are inadequate compared with nonrural schools. As I stated previously, all of these conclusions are provisional. They are not the result of an exhaustive analysis of Kentucky test score trends, but are intended instead to present the rudiments of how such an analysis would proceed. A more complete analysis would examine all of the Kentucky test results that are available for the different grade levels in science and math, and it would include other characteristics of schools and of their student populations in order to diagnose more comprehensively the reasons for variation in test score trends. Only a comprehensive approach such as this would present policymakers with the needed wide-ranging information that could guide the fashioning of new interventions and programs.

Growth modeling is a flexible tool for use in education policy research. This brief presentation has shown some basic ways in which this technique may be used to address policy concerns regarding math and science achievement trends. Because The No Child Left Behind Act mandates that U.S. public schools must demonstrate "adequate yearly progress" toward the goal of 100 percent proficiency, it is likely that growth modeling will become the policy research tool of choice when studying multiyear trends in school and student achievement. Another use of growth modeling that bears a resemblance to the examples presented in this paper is that of testing the effectiveness of a policy intervention. In such a test, the policy intervention may be measured either categorically or with a continuous scale. The intervention can also be either time-

invariant, as rural location was in my examples above, or time-dependent, as the percent of low-income students was. An example of a time-dependent policy intervention could be one in which different stages of intervention are allowed to occur at different time intervals for various target organizations or populations. The procedures described in the examples that were presented above, including the tests for a regression artifact and spuriousness, would remain essentially the same.

Note

This paper was presented at the 12th Annual International Learning Conference held at the University of Granada, Granada, Spain during July 2005. The present version of the paper has benefited from comments by Debbie Abell, Louise Cooper, and Jesse Lowe. Any errors that remain are the author's responsibility.

References

Campbell, D. T., & Kenny, D. A. (1999). *A primer on regression artifacts*. New York: Guilford Press.

Education Alliance. (2004). *Closing the achievement gap: A focus on rural schools*. Charleston, WV: Author. Retrieved May 18, 2005 from http://www.educationalliance.org/RuralSchools.pdf.

Fan, X., & Chen, M. J. (1999). Academic achievement of rural school students: A multi-year comparison with their peers in suburban and urban schools. *Journal of Research in Rural Education, 15,* 31-46.

Greenberg, E. J., & Teixeira, R. A. (1995). Nonmetro student achievement on par with metro. *Rural Development Perspectives, 10,* 17-23.

Israel, G. D., Beaulieu, L. J., & Hartless, G. (2001). The influence of family and community social capital on educational achievement. *Rural Sociology, 66, 43-68.*

Khattri, N., Riley, K. W., & Kane, B. (1997). Students at risk in poor, rural areas: A review of the research. *Journal of Research in Rural Education, 13,* 79-100.

Ladd, H. F. (Ed.). (1996). *Holding schools accountable: Performance-based reform in education*. Washington, DC: The Brookings Institution.

Luke, D. A. (2004). *Multilevel modeling* (Quantitative Applications in the Social Sciences, 143). Thousand Oaks, CA: Sage.

National Research Council. (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*, edited by J. Pellegrino, N. Chudowsky, and R. Glaser. Board on Testing and Assessment, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis 26*, 237-257.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Reeves, E. B. (2000). High stakes accountability and contextual effects: An empirical study of the fairness issue. *Research in the Schools, 7,* 49-58.

_____. (2003). Disentangling the effects of nonmetro location and student poverty on school performance/improvement: Implications for equitable excellence in Kentucky public schools. *Journal of Research in Rural Education, 18,* 17-30.

Reeves, E. B., & Bylund, R. A. (2005). Are rural schools inferior to urban schools? A multilevel analysis of school accountability trends in Kentucky. *Rural Sociology, 70*, 360-386.

Roscigno, V. J. & Crowley, M. L. (2001). Rurality, institutional disadvantage, and achievement/attainment. *Rural Sociology, 66*, 268-292.

Seltzer, M., Choi, K., & Thum, Y. M. (2003). Examining relationships between where students start and how rapidly they progress: Using new developments in growth modeling to gain insight into the distribution of achievement within schools. *Educational Analysis and Public Policy, 25*, 263-286.

Singer, J. D. & Wilbert, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

United States Government Accountability Office. (2004, September). No child left behind act: Additional assistance and research on effective strategies would help small rural districts (GAO-04-909). Retrieved February 15, 2005 from www.gao.gov/new.items/ d04909.pdf.

Webster, B. J., & Fisher, D. L. (2000). Accounting for variation in science and mathematics achievement: A multilevel analysis of Australian data. School *Effectiveness and School Improvement,11*, 339-360.

Williams, J. H. (2005). Cross-national variations in rural mathematics achievement: A descriptive overview. *Journal of Research in Rural Education, 20*, 1-18. Retrieved April 27, 2005 from www.umaine.edu/jrre/20-5.htm.

Young, D. (1998). Rural and urban differences in student achievement in science and mathematics: A multilevel analysis. *School Effectiveness and School Improvement, 9*, 386-418.

_____. (2000). Rural differences in student achievement: The effect of student perceptions. *Educational Research and Evaluation, 6*, 207-228.